This paper appears in *Environmetrics*, 2016, DOI: 10.1002/env.2389 (http://dx.doi.org/10.1002/env.2389).
The document below represents the final revision (preprint) submitted to the journal prior to publication.

Spatial Modelling Framework for the Characterisation of Rainfall Extremes at Different Durations and under Climate Change

Eric A. Lehmann[1], Aloke Phatak[2], Alec Stephenson[3], Rex Lau[4]

[1]CSIRO Data61, Acton ACT, Australia
[2]Department of Mathematics and Statistics, Curtin University, Bentley WA, Australia
[3]CSIRO Data61, Clayton South VIC, Australia
[4]CSIRO Data61, Floreat WA, Australia

Abstract

This paper describes a statistical modelling framework for the characterisation of rainfall extremes over a region of interest. Using a Bayesian hierarchical approach, the data are assumed to follow the generalised extreme value distribution, whose parameters are modelled as spatial Gaussian processes in the latent process layer. We also integrate a parametric relationship between precipitation maxima accumulated over increasing durations. Inference of the model parameters is thus improved by pooling information across both space and accumulation duration. In addition, we propose and investigate two different approaches for the integration of daily and sub-daily rainfall data within the framework. We also demonstrate how information from a regional climate model can be integrated to enable the investigation of future projections of extreme rainfall characteristics.

We apply the proposed methodology to precipitation datasets from two large-scale study regions located on the east coast of Australia. The models are fitted using Markov chain Monte Carlo simulations, and we present estimated model parameters and posterior inferences of return levels at various durations and sites of interest. We demonstrate the effectiveness of the framework in spatially extrapolating the inference to locations other than those at which direct rainfall measurements are available. We also provide comparisons between rainfall extremes at various durations obtained for the current climate and those based on future projections from a regional climate model. Both methods proposed for the integration of daily and sub-daily records were found to yield similar results in terms of model performance and computational requirements.

Keywords—Bayesian hierarchical modelling, extreme value theory, Hershfield factor, extremal index, Markov chain Monte Carlo, intensity–frequency–duration curve.

1 Introduction

Understanding the relationship between rainfall intensity, frequency, and duration is important for the design of hydraulic infrastructure that can withstand extreme rainfall events (Koutsoyiannis *et al.*, 1998). As a result of the changing nature of extreme rainfall (IPCC, 2013), however, intensity–frequency–duration (IFD) relationships estimated using current climate observations are unlikely to be suitable for infrastructure design in the future (Cheng and AghaKouchak, 2014). In this paper, we propose a Bayesian hierarchical model (BHM) that pools information across space to produce estimates of IFD curves and their associated uncertainty. This framework is novel in that it allows the integration, in a spatial dependence context, of two disparate rainfall datasets from pluviometer (5-minute duration) and daily stations. The

combination of such datasets was already explored by Muller *et al.* (2008) and Koutsoyiannis *et al.* (1998), although only on a site-by-site basis. In order to model IFD curves, which require annual maxima of sub-daily rainfall accumulated over increasing durations, we further incorporate a duration dependence relationship due to Koutsoyiannis *et al.* (1998) into the spatial BHM, which improves the estimation of model parameters by combining information available across durations.

In the statistical literature, Bayesian hierarchical models have been widely used to model rainfall extremes of single—usually daily—duration, e.g. (Atyeo and Walshaw, 2012; Davison *et al.*, 2012; Dyrrdal *et al.*, 2015; Ghosh and Mallick, 2011; Geirsson *et al.*, 2015). Little has been done, however, on modelling IFD curves, with the exception of the works by Muller *et al.* (2008) and Van de Vyver (2015), who, for a single station only, also integrate the duration relationship of Koutsoyiannis *et al.* (1998). Wang and So (2016) use a different approach altogether where the duration variable is treated as a "proximity" measure included as a covariate in the latent process structure. Outside the Bayesian context, IFD curves are usually produced by regional frequency analysis (RFA) (Green *et al.*, 2012; Hosking and Wallis, 2005). RFA combines data from stations into regions with similar characteristics, such as geographical location, in order to increase the sample size and hence the precision of extreme rainfall estimates. Modelling is carried out on a duration-by-duration basis, and there is currently no methodology for quantifying uncertainty at all stages of RFA, or for the integration of climate variables to assess the impact of climate change. The spatial BHM that we propose here can be seen as an alternative to RFA. The spatial Gaussian process introduced in the process layer of the BHM is the analogue of the regionalisation described above. On the other hand, the BHM allows us to combine information from different durations, with uncertainty estimates arising naturally from the estimation framework. We also use an innovative approach to integrate the outputs from a regional climate model (RCM) into the spatial BHM in order to estimate potential changes in IFD curves resulting from various climate change scenarios.

The impact of climate change on IFD curves is commonly assessed by comparing the change in return levels between a baseline period and a future period, assuming that the extremes are locally stationary. Srivastav *et al.* (2014) categorize approaches to using information from global climate models (GCMs) and RCMs to assess the impact of climate change on IFD curves into one of two classes, and provide recent citations in each category. The approaches include: (a) change factor or 'delta' methods that use only information from G/RCM outputs in a baseline period and a future period to estimate the impact of climate change on return levels across a range of durations; and (b) bias-correction methods, in which differences between observed data and G/RCM outputs for a baseline period are then used to modify the G/RCM outputs in future periods, from which IFD curves are then calculated. Applications of these approaches have been limited to individual stations (Peck *et al.*, 2012; Yilmaz *et al.*, 2014). By contrast, the method we describe here represents a spatial extremes model that incorporates information about climate change by relating observational data and climate model outputs at the level of the parameters of the generalized extreme value distribution that is used to model extreme rainfall at different durations.

In summary, the main contributions made by this paper are encapsulated by the design of a flexible and unified framework for the calculation of IFD curves that simultaneously:

(a) combines daily and sub-daily datasets of rainfall maxima, with potentially missing data
(b) achieves improved inference of model parameters by drawing on information from surrounding stations, pooling data across durations, and making use of the r largest rainfall maxima
(c) provides estimates of IFD curves and return levels at ungauged locations
(d) is able to incorporate a variety of covariates, including climate variables that enable future

projections of IFD curves, and

(e) enables the propagation of error through the model, thus providing straightforward estimates of uncertainty for all quantities of interest.

The proposed method thus improves on existing and more traditional methods for the calculation of IFD curves and return levels, which either:

(a) rely on multi-stage approaches, e.g. for the calculation of IFD curves (Koutsoyiannis *et al.*, 1998; Hosking and Wallis, 2005) or for the investigation of climate change (Srivastav *et al.*, 2014), thereby complicating the process of error characterisation

(b) are limited to single durations, e.g. (Geirsson *et al.*, 2015; Dyrrdal *et al.*, 2015; Bracken *et al.*, 2016), or operate on a duration-by-duration basis (Green *et al.*, 2012; Yilmaz *et al.*, 2014)

(c) are limited to single stations, e.g. (Muller *et al.*, 2008; Van de Vyver, 2015; Peck *et al.*, 2012), or

(d) only account for one type (daily or sub-daily) of precipitation data, e.g. (Wang and So, 2016; Cooley *et al.*, 2007; Lehmann *et al.*, 2013; Davison *et al.*, 2012).

Naturally, the applicability of the proposed approach is generic in the sense that it can be applied to any region where the relevant datasets are available. In this work, we demonstrate the wide-scale and operational feasibility of our methodology by applying it to two large regions in Australia, covering up to 580,000 km^2 and containing upwards of 1300 stations, which are typically one to two orders of magnitude larger than most other existing case studies on spatial rainfall extremes, e.g. (Atyeo and Walshaw, 2012; Wang and So, 2016; Van de Vyver, 2012; Cooley *et al.*, 2007).

The manuscript is organised as follows. Sections 2–5 describe the structure of our approach, integration of pluviometer and daily data, and integration of climate change information. In Sec. 6, we describe the data that were used and present some simulation results, and in Sec. 7 we discuss extensions and improvements that could be made to the proposed model.

2 Modelling of precipitation maxima

2.1 Extreme value distribution

The generalised extreme value (GEV) distribution is often used to model rainfall extremes (Coles, 2001). Under certain conditions, the distribution of annual maximum rainfall Y of a given duration can be modelled by a GEV distribution with location $\mu \in (-\infty, +\infty)$, scale $\sigma > 0$, and shape $\xi \in (-\infty, +\infty)$. Generalising this formulation to model the r largest annual rainfall maxima $\mathbf{y} = [y^{(1)}, \ldots, y^{(r)}]^T$, with $r \geq 1$ and $y^{(1)} \geq y^{(2)} \geq \ldots \geq y^{(r)}$, leads to the r-largest order GEV density function (Coles, 2001):

$$\mathrm{GEV}_r\left(\mathbf{y}; \mu, \sigma, \xi\right) = \exp\left\{-\left[1+\xi\cdot\left(\frac{y^{(r)}-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}\cdot\prod_{i=1}^{r}\frac{1}{\sigma}\left[1+\xi\cdot\left(\frac{y^{(i)}-\mu}{\sigma}\right)\right]^{-1-1/\xi} \quad (1)$$

where $1+\xi\cdot(y^{(r)}-\mu)/\sigma > 0$. Making use of $r > 1$ rainfall maxima leads to improved model estimates by essentially increasing the amount of data available for inference.

Extreme quantiles of the GEV distribution (Coles, 2001) can be calculated to produce intensity–frequency–duration (IFD) curves. At a given location, the IFD curve describes the rainfall intensity (mm/h) that can be expected over some accumulation duration and for a given return period, or alternatively, for a given annual exceedance probability (sometimes also

described as a "1–in–N years" event). These curves represent an important source of information for the design of infrastructure, and to assess the social and economic impacts resulting from extreme rainfall events (Pilgrim, 1997).

2.2 Dependence model across durations

Koutsoyiannis *et al.* (1998) re-parameterise the location parameter as $\tilde{\mu} = \mu/\sigma$ and demonstrate that both $\tilde{\mu}$ and the shape parameter ξ can be assumed to remain constant for extreme rainfall data accumulated over various durations. On the other hand, the scale parameter is dependent on the accumulation duration d, and the following relationship is used in Koutsoyiannis *et al.* (1998) to model this dependence:

$$\tilde{\sigma} = \tilde{\sigma}(d) = \frac{\sigma \cdot d}{(d+\theta)^{\eta}}, \quad \theta \geq 0, \quad \eta \in [0,1], \tag{2}$$

where $\tilde{\sigma}$ is the scale parameter of extreme precipitation accumulated over duration d, and σ is now a duration-independent parameter. In Eq. (2), the duration d in the numerator is required when rainfall is expressed in terms of depths (mm) rather than intensities (mm/h). In addition to the three standard GEV parameters μ, σ and ξ, this formulation thus leads to two additional model parameters, namely the offset term θ and exponent η.

The relationship in Eq. (2) was derived empirically by Koutsoyiannis *et al.* (1998) through analyses of several datasets. As the authors write, an expression of this form "encapsulates the experience from several IFD studies," and it has since also been used successfully with other datasets (Van de Vyver, 2015; Muller *et al.*, 2008; Peck *et al.*, 2012). Other empirical IFD relationships were investigated by Garcia-Bartual and Schneider (2001). Extensive exploratory data analyses were performed on our precipitation datasets to assess the applicability of the relationship in Eq. (2), which appears to fit the data well. Sec. S3.1 in the accompanying Supplementary Material document presents and further discusses the results of this analysis.

In contrast to traditional methods operating on a duration-by-duration basis (e.g. RFA), making use of this dependence relationship allows us to achieve improved estimates of the model parameters by effectively pooling the precipitation data across durations. We also note in passing that Nadarajah *et al.* (1998) showed that the ordering of rainfall maximum depths at increasing durations implies restrictions on the parameters of the extreme value distributions. In particular, when $\tilde{\mu}$ and ξ are roughly constant across all durations, the scale parameters $\tilde{\sigma}(d)$ are subject to the order restriction $\tilde{\sigma}(d_1) \leq \tilde{\sigma}(d_2)$ when $d_1 < d_2$, which is indeed enforced by the empirical expression in Eq. (2).

3 Bayesian hierarchical model

Here, we present a statistical model for the estimation of GEV parameters from rainfall data accumulated over different durations (pluviometer data). We introduce the formulation $\tilde{\mu} = \mu/\sigma$ and the dependence model of Eq. (2) into this framework, which we then extend in Sec. 4 to integrate daily precipitation measurements.

Our approach is based on a BHM (Banerjee *et al.*, 2003; Stephenson, 2016), with the data and process layers described in the following. Technical details regarding the prior distributions, Bayesian inference and conditional distributions can be found in Sections S1 and S5 of the Supplementary Material. The use of a Bayesian approach for the calculation of IFD curves improves on existing (traditionally multi-step) methods by accounting for the spatial and duration-related dependences, error characterisation, integration of covariates, as well as parameter inference and extrapolation, within a single unified framework.

3.1 Data model

The r largest annual precipitation maxima $\mathbf{y}_{s,t,i}$, recorded at station $s \in \{1, \ldots, S\}$ in year $t \in \{1, \ldots, T_s\}$ and for an accumulation duration $d_{s,i}$, $i \in \{1, \ldots, D_s\}$, are assumed to be independent conditional on the station-specific GEV parameters:

$$\mathbf{y}_{s,t,i} \sim \mathrm{GEV}_r^*(\tilde{\mu}_s, \sigma_s, \xi_s, \theta_s, \eta_s) \tag{3}$$

$$= \mathrm{GEV}_r(\tilde{\mu}_s, \tilde{\sigma}_{s,i}, \xi_s) \tag{4}$$

where the station-specific, duration-independent parameters σ_s, θ_s and η_s are used to calculate the duration-specific scale parameter $\tilde{\sigma}_{s,i}(d_{s,i})$ using Eq. (2).

Our formulation assumes stationarity over the period during which the data are recorded: at each station, the rainfall maxima are pooled together into a single dataset, irrespective of the acquisition year. Consequently, the model can incorporate stations with various levels of missing data, i.e., stations with differing sets of acquisition years and varying record lengths T_s. Similarly, the relationship in Eq. (2) allows each station to have different sets of accumulation durations with varying cardinality D_s. Let $\tilde{\boldsymbol{\mu}} = [\tilde{\mu}_1, \ldots, \tilde{\mu}_S]^T$, with $\boldsymbol{\sigma}$, $\boldsymbol{\xi}$, $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ defined similarly, and let $\mathbf{Y} = \{\mathbf{y}_{s,t,i}\}$ represent the set of all available precipitation maxima. The composite likelihood for the extreme rainfall data is thus given by

$$p(\mathbf{Y}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}) = \prod_{s=1}^{S} \prod_{t=1}^{T_s} \prod_{i=1}^{D_s} \mathrm{GEV}_r^*(\mathbf{y}_{s,t,i}; \tilde{\mu}_s, \sigma_s, \xi_s, \theta_s, \eta_s). \tag{5}$$

3.2 Process model

We assume that the GEV parameters vary smoothly over space, and thus model these variables as the addition of a linear regression on covariates and a spatial process. For each $\chi \in \{\tilde{\mu}, \sigma, \xi, \theta, \eta\}$, we use

$$g_\chi(\boldsymbol{\chi}) = \mathbf{X}_\chi \boldsymbol{\beta}_\chi + \mathcal{P}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi) \tag{6}$$

where $\mathbf{X}_\chi$ represents the S-by-$(N_\chi + 1)$ design matrix of N_χ (scaled and centered) covariates with intercept term at each station, and $\boldsymbol{\beta}_\chi$ is the corresponding vector of $N_\chi + 1$ regression coefficients. The (componentwise) link function $g_\chi(\cdot)$ is chosen as the identity for $\tilde{\boldsymbol{\mu}}$ and $\boldsymbol{\xi}$, $\log(\cdot)$ for $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$, and $\mathrm{logit}(\cdot)$ for $\boldsymbol{\eta}$.

In Eq. (6), $\mathcal{P}_\chi(\cdot)$ represents a spatially correlated, zero-mean Gaussian random process (multivariate normal) with covariance matrix $\boldsymbol{\Sigma}_\chi$ defined on the set of station locations $\boldsymbol{\ell} = \{\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_S\}$. Different approaches can be used to model the spatial dependence of the parameters (Davison *et al.*, 2012). Here we make use of the parametric family of (isotropic) exponential correlation functions to model the smoothness of the parameters over space. The (i, j)-th entry of the covariance matrix is thus defined as:

$$[\boldsymbol{\Sigma}_\chi]_{(i,j)} = \alpha_\chi \cdot \exp\left(-\frac{\|\boldsymbol{\ell}_i - \boldsymbol{\ell}_j\|}{\lambda_\chi}\right), \quad i, j = 1, \ldots, S \tag{7}$$

with α_χ and λ_χ denoting the sill and range of the covariance function, respectively, and $\boldsymbol{\ell}_i$ denoting the location of the i-th station. This spatial model essentially allows to: (a) improve the model inference by drawing on data available at surrounding stations (akin to regionalisation in RFA), and (b) provide estimates of GEV parameters at ungauged locations, where no direct rainfall measurements are available.

4 Integration of daily data

The BHM presented in Sec. 3 represents a general framework to model extreme precipitation data accumulated over a range of durations. In addition to pluviometer data (e.g. registered over 5 min intervals), records of daily rainfall usually also exist in a region of interest. Daily rainfall is measured in a single 24 h period, typically from 9:00 a.m. on one day to 9:00 a.m. on the next. These measurements usually extend further back in time and originate from a denser network of rain gauges, and thus provide valuable additional information on extreme precipitation (Hershfield, 1961; Koutsoyiannis *et al.*, 1998; Muller *et al.*, 2008).

In this section, we propose a generic approach to combine, and thus take advantage of, both sources of information in a unified BHM framework. We first present a re-parametrisation of the scale parameter that enables the combination of such datasets. In Sections 4.2 and 4.3, we then describe two different approaches to integrating daily rainfall data into the BHM. Each method provides an alternative solution to the problem of estimating annual maximum rainfall in *any* 24 h period from annual daily maxima.

4.1 Re-parameterisation for 24 hour data

To account for pluviometer data accumulated over various durations, the model presented in Sec. 3 incorporates the duration offset and exponent terms θ and η, respectively. Direct inference of these parameters using precipitation data from daily stations is impossible, as these data correspond to a single accumulation duration of 24 hours (recorded at 9:00 a.m. each day). We thus propose a re-parameterisation of the scale parameter σ to account for this specificity.

In addition to the set of pluviometer stations, $s \in \{1, \ldots, S\}$, let us assume that rainfall data from a set of U daily stations are available: $u \in \{S+1, \ldots, S+U\}$. In order to integrate daily precipitation data into the current modelling framework, we express the GEV scale parameter at station n (pluviometer or daily) in terms of $\tilde{\sigma}_n^{(24)}$, i.e. the scale parameter $\tilde{\sigma}_{n,i}$ for duration $d_{n,i} = 24$ h. Simple algebraic manipulations of Eq. (2) show that the (duration-specific) scale parameter for pluviometer station $s \in \{1, \ldots, S\}$ at duration $d_{s,i}$ then becomes:

$$\tilde{\sigma}_{s,i} = \frac{d_{s,i} \cdot \tilde{\sigma}_s^{(24)}}{24} \cdot \left(\frac{24 + \theta_s}{d_{s,i} + \theta_s} \right)^{\eta_s} . \tag{8}$$

This definition of the scale coefficient allows for a direct use of $\tilde{\sigma}_u^{(24)}$ for the daily stations $u \in \{S+1, \ldots, S+U\}$, thereby avoiding the need to model the (pluviometer-specific) parameters θ and η at those stations. In the following, only three GEV parameters $(\tilde{\mu}_u, \tilde{\sigma}_u^{(24)}, \xi_u)$ are thus modelled at each daily station u, while the pluviometer stations $s \in \{1, \ldots, S\}$ still have five as before, namely $(\tilde{\mu}_s, \tilde{\sigma}_s^{(24)}, \xi_s, \theta_s, \eta_s)$.

4.2 Hershfield factor

Let us denote by $\mathbf{y}_{u,t}^{(\mathrm{D})}$ the annual maxima of daily rainfall from daily station $u \in \{S+1, \ldots, S+U\}$ and year $t \in \{1, \ldots, T_u\}$. A standard technique used in quantitative hydrology is to use the multiplicative Hershfield factor (HF) H to convert annual maxima of daily rainfall into annual maxima of 24 h rainfall (Hershfield and Wilson, 1958; van Montfort, 1997). In previous studies (Hershfield, 1961), this factor was found to be $H \approx 1.13$ for precipitation data.

The first approach to integrating daily data within the BHM thus consists in using the model described so far in conjunction with a "corrected" set of daily precipitation maxima:

$$\mathbf{Y} = \mathbf{Y}^{(\mathrm{P})} \cup \left(H \cdot \mathbf{Y}^{(\mathrm{D})} \right) = \{\mathbf{y}_{s,t,i}\} \cup \left\{ H \cdot \mathbf{y}_{u,t}^{(\mathrm{D})} \right\} \tag{9}$$

where the daily maxima $\mathbf{y}_{u,t}^{(\mathrm{D})}$ have been transformed into "pluviometer-equivalent" maxima corresponding to (a single) accumulation duration $d_u = 24$ h, $u \in \{S+1, \ldots, S+U\}$. The composite likelihood calculations in Eq. (5) will consequently only include one duration term for each daily station u ($D_u = 1$):

$$p\left(\mathbf{Y}|\tilde{\boldsymbol{\mu}}_+, \tilde{\boldsymbol{\sigma}}_+^{(24)}, \boldsymbol{\xi}_+, \boldsymbol{\theta}, \boldsymbol{\eta}\right) = \prod_{s=1}^{S}\prod_{t=1}^{T_s}\prod_{i=1}^{D_s} \mathrm{GEV}_r^*\left(\mathbf{y}_{s,t,i}; \tilde{\mu}_s, \tilde{\sigma}_s^{(24)}, \xi_s, \theta_s, \eta_s\right) \times \prod_{u=S+1}^{S+U}\prod_{t=1}^{T_u} H^r \cdot \mathrm{GEV}_r\left(H \cdot \mathbf{y}_{u,t}^{(\mathrm{D})}; \tilde{\mu}_u, \tilde{\sigma}_u^{(24)}, \xi_u\right) \quad (10)$$

with the notation $\tilde{\boldsymbol{\sigma}}_+^{(24)} = [\tilde{\sigma}_1^{(24)}, \ldots, \tilde{\sigma}_{S+U}^{(24)}]^T$ denoting the combined vector of daily and pluviometer parameters (similar definitions for $\tilde{\boldsymbol{\mu}}_+$ and $\boldsymbol{\xi}_+$). The factor H^r in the second term of Eq. (10) represents the Jacobian of the transformation applied to the daily maxima, and allows a direct comparison of the model selection results presented further below in Sec. 6.2 (information criteria, see also Sec. S4 in the Supplementary Material).

One advantage of this approach is that the same model definitions can be used for both the daily and pluviometer stations. Using Eq. (8), the parameters $\tilde{\sigma}_s^{(24)}$, θ_s and η_s in the first term of Eq. (10) are used to compute the duration-specific scale parameter $\tilde{\sigma}_{s,i}$ for pluviometer stations. In what follows, this model will be referred to as the BHM-HF approach.

4.3 Extremal index

The second approach uses the extremal index (EI) to link the GEV density of the daily data to that of an equivalent set of pluviometer maxima recorded over a 24 h accumulation duration. Results in Robinson and Tawn (2000) show that, under specific conditions and assumptions, the daily maxima have a GEV distribution given by:

$$\mathbf{y}_{u,t}^{(\mathrm{D})} \sim \mathrm{GEV}_r\left(\tilde{\mu}_u^{(\mathrm{D})}, \sigma_u^{(\mathrm{D})}, \xi_u^{(\mathrm{D})}\right) \quad (11)$$

where the "daily-equivalent" GEV parameters can be expressed in terms of the model parameters as follows:

$$\tilde{\mu}_u^{(\mathrm{D})} = \tilde{\mu}_u \cdot \Lambda^{\xi_u} - \xi_u^{-1} \cdot \left(\Lambda^{\xi_u} - 1\right) \quad (12)$$

$$\sigma_u^{(\mathrm{D})} = \tilde{\sigma}_u^{(24)} \cdot \Lambda^{-\xi_u} \quad (13)$$

$$\xi_u^{(\mathrm{D})} = \xi_u \quad (14)$$

with

$$\Lambda = \frac{\delta \cdot \left[1 - (1 - \Theta_\delta)^{1/\delta}\right]}{\Theta_\delta}. \quad (15)$$

The parameter Θ_δ is known as the extremal index associated with a rainfall process observed at δ^{-1} times the original (fundamental) rate. For instance, by comparison to a time series of 24 h accumulations computed from pluviometer data recorded at 5 min intervals, a time series of daily rainfall can be seen as having been sub-sampled by a factor $\delta = (24 \cdot 60)/5 = 288$. In this work, we use a value of $\Theta_\delta = 0.9$, which was obtained by using the empirical approach of Van de Vyver (2012).

To use the extremal index, we rewrite the composite likelihood in Eq. (5) as:

$$p\left(\mathbf{Y}|\tilde{\boldsymbol{\mu}}_+, \tilde{\boldsymbol{\sigma}}_+^{(24)}, \boldsymbol{\xi}_+, \boldsymbol{\theta}, \boldsymbol{\eta}\right) = \prod_{s=1}^{S}\prod_{t=1}^{T_s}\prod_{i=1}^{D_s} \mathrm{GEV}_r^*\left(\mathbf{y}_{s,t,i}; \tilde{\mu}_s, \tilde{\sigma}_s^{(24)}, \xi_s, \theta_s, \eta_s\right) \times \prod_{u=S+1}^{S+U}\prod_{t=1}^{T_u} \mathrm{GEV}_r\left(\mathbf{y}_{u,t}^{(\mathrm{D})}; \tilde{\mu}_u^{(\mathrm{D})}, \sigma_u^{(\mathrm{D})}, \xi_u^{(\mathrm{D})}\right). \quad (16)$$

This model is subsequently referred to as the BHM-EI approach.

5 Integration of climate change information

In the process model of Sec. 3.2, the Gaussian processes $\mathcal{P}_{(\cdot)}$ are used to model the spatially correlated (local) variability of the GEV parameters in the vicinity of each station. By contrast, the covariates contained in the design matrices $\mathbf{X}_{(\cdot)}$ account for the large-scale variability across the whole study region. Following this line of thought, we use the output from a regional climate model (RCM) as a climatological covariate to account for the effects of climate in the computation of return levels and IFD curves. This approach then allows us to investigate the impact of climate change on extreme rainfall levels by calculating future projections of IDF curves.

Because RCM outputs already account for geographical variables such as elevation and location, we exclude such covariates from the process layer in Eq. (6), which then becomes:

$$g_\chi(\boldsymbol{\chi}) = \beta_{\chi,0} \cdot \mathbf{1} + \beta_{\chi,1} \cdot g_\chi(\boldsymbol{\chi}_{\mathrm{RCM}}) + \mathcal{P}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi) \quad (17)$$

where $\boldsymbol{\chi}_{\mathrm{RCM}}$ is the vector of GEV parameters obtained from the RCM output at the grid cells containing each station (nearest RCM grid point). This nearest-neighbour approach leads to reasonable results given the smoothness of the RCM surfaces used in our study. If necessary, a more general approach could be used where the climatological covariate at each site is spatially interpolated on the basis of the neighbouring grid cells.

The proposed BHM in Eq. (17) is fitted on the basis of the available (pluviometer and daily) precipitation data and the RCM parameters obtained over the same time period. To assess how precipitation extremes might change in the future, model-based estimates of the GEV parameters for a future climate (at gauged or ungauged locations) can then be inferred by temporally extrapolating the BHM results using the corresponding RCM outputs obtained for the future climate. This allows us to generate future-climate IFD curves and to compare them to those from the baseline period. Also, when treated as a source of spatial information, integration of G/RCM variables in the process model of the BHM has the side effect of strengthening the spatial extrapolation of results in regions where the station density is low in comparison with the resolution of the G/RCM grid (Geirsson *et al.*, 2015).

Note that incorporating parameter estimates from an RCM into the process layer also allows for an implicit evaluation of how well the RCM reproduces observed extremes in the baseline period. If it does so well, we can expect the marginal posterior probabilities of RCM-related regression coefficients ($\beta_{\chi,1}$ in Eq. (17)) to be small near zero. In other words, there would be some relationship between the parameter estimates from the data and those from the RCM over the same time period.

6 Results

6.1 Preliminaries

The methods described in this paper can be used to derive a large number of different (possibly nested) model variants. For both the BHM-HF and BHM-EI approaches, a number of different covariates such as latitude/longitude, elevation, slope orientation, RCM outputs, and so on, can be combined (included or excluded) in various ways at the process level for each GEV parameter. For instance, earlier works such as Koutsoyiannis *et al.* (1998) and Davison *et al.* (2012) model the shape parameter ξ as a constant over space, while the process models for the GEV parameters μ and σ are defined as functions of spatial coordinates. Because it would be difficult to compare a large number of different model variants in this work, we focus specifically on the following analyses. In Sec. 6.2, we compare the two proposed models based on the same process layer definitions using standard geographical covariates. On the basis of these results, the BHM-HF model is used to demonstrate the integration of climate change information in Sec. 6.3 (current vs. future IFD curves). Finally, Sec. 6.4 investigates the effects of adding daily stations to the dataset of pluviometer records.

The consideration of two or more models for precipitation extremes inevitably leads to the question of whether one is preferable in practice. We address this issue by using standard measures of model fit including posterior predictive checking, posterior coverage and predictive accuracy, as well as the deviance information criterion (DIC) and widely applicable information criterion (WAIC). The derivation of these model selection and validation parameters for the proposed models is provided in Sec. S4 of the Supplementary Material document.

All the results presented in this sections were obtained on the basis of 150,000 iterations of the MCMC algorithms, with the first 10,000 iterations discarded (burn-in), and thinned by a factor 40. The convergence of the algorithms was assessed through visual inspections of the MCMC chains (trace plots), as well as standard diagnostics for Markov chains (Robert and Casella, 2010; Brooks and Roberts, 1998). Good mixing properties and no identifiability issues were observed for any of the simulated model parameters. Further aspects of the computational implementation and convergence diagnostics are discussed in Sec. S2 of the Supplement.

6.2 Comparison of BHM-HF and BHM-EI

For a comparative assessment of the proposed methods, both models are defined as described in Sections 3 and 4, with the process models for all GEV parameters in Eq. (6) defined on the basis of: (a) an intercept, (b) spatial coordinates (eastings and northings in km), and (c) elevation (height above sea level).

Precipitation data from two large-scale regions in Australia are used in the evaluation. The first dataset (Fig. 1, top) contains rainfall data in the Greater Sydney region (SYD) from 872 stations (covering approx. 156,000 km^2), with $S_{\text{SYD}} = 310$ pluviometer and $U_{\text{SYD}} = 562$ daily stations. The second dataset (Fig. 1, bottom), covering approx. 580,000 km^2, contains 1348 stations located in the South East Queensland (SEQ) region: $S_{\text{SEQ}} = 240$ pluviometer and $U_{\text{SEQ}} = 1108$ daily stations. Both datasets were subjected to a thorough quality assurance process, including a number of different exploratory analyses. The pluviometer data at 5 min intervals were subsequently accumulated over 12 different durations, namely 5, 10, 15 and 30 minutes, and 1, 2, 3, 6, 12, 24, 48 and 72 hours. All stations within these datasets provide at least eight years' worth of annual top-two precipitation maxima ($r = 2$) during the period from 1961 to 2000. For all durations, the second maximum was selected at least three days before or after the top maximum to ensure independence. To calculate the predictive accuracy measure, randomly selected pluviometer stations were set aside in both regions: 50 in the SYD region

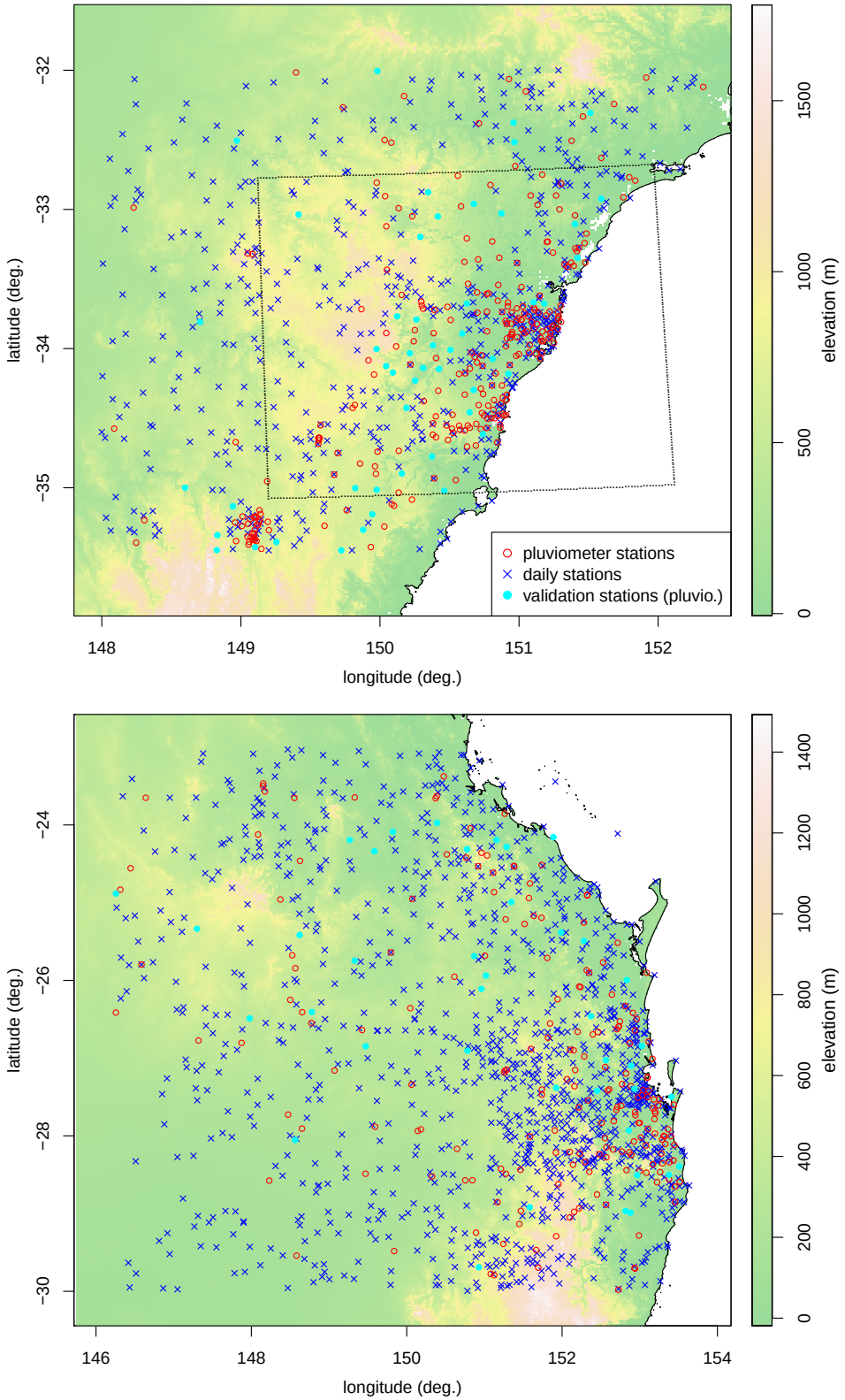

Figure 1: Datasets of daily and pluviometer stations in the Greater Sydney region (SYD, top) and the South East Queensland region (SEQ, bottom). The dotted lines in the SYD plot show the extent of the RCM grid used for the assessment of climate impact on IFD curves.

	SYD		SEQ	
	BHM-HF	BHM-EI	BHM-HF	BHM-EI
posterior predictive check	**0.355**	0.351	**0.355**	0.350
posterior coverage (%)	94.37	**94.40**	**94.92**	94.83
predictive accuracy	**4.60**	3.95	**2.51**	2.10
DIC	**263.1**	278.8	**1106.3**	1314.6
WAIC	153.3	**144.9**	**928.5**	1107.2

Table 1: Model selection and validation results. For ease of interpretation, predictive accuracy scores were scaled by a factor 10^3, while fixed constants were removed from the DIC and WAIC results ($1069 \cdot 10^3$ for SYD, $1384 \cdot 10^3$ for SEQ). Boldface values indicate better results.

and 40 in the SEQ region (see Fig. 1). For all other validation criteria, the full precipitation datasets were used, i.e., validation stations included.

Table 1 shows the model selection and validation results for the BHM-HF and BHM-EI approaches in both geographical regions. The posterior predictive checks yield similar results for both algorithms, and do not point to any systematic issues that might be symptomatic of potential model failings. The posterior coverage results do not make a clear-cut distinction between BHM-HF and BHM-EI either, indicating that both models describe the data well. Further insight into how well the two models fit the data at various durations is provided in Sec. S3.2 in the Supplement. The predictive accuracy scores, as well as the DIC and WAIC measures for the SEQ region, indicate that BHM-HF might perform somewhat better than BHM-EI.

Fig. S3 in Sec. S3.3 of the Supplementary Material contains plots of the posterior densities of the main model parameters, showing that both approaches yield similar results. The calculation of IFD curves is mainly driven by $\tilde{\sigma}^{(24)}$, and in this respect, Fig. S3 shows that results from both the BHM-HF and BHM-EI models are in relative agreement (a further discussion is provided in Sec. S3.3). A detailed investigation (not shown here) of the IFD curves at gauged and ungauged locations shows that return levels generated by BHM-EI are on average about 4% larger than those from BHM-HF. The width of the uncertainty band around the IFD curves also varies between algorithms, with the BHM-EI uncertainty in general larger for small return periods (e.g. 10 years), and smaller for large return periods (e.g. 500 years).

The computational requirements of both approaches are similar (see Table S3 in the Supplement), with the overall performance varying slightly depending on the number and type (daily vs. pluviometer) of stations in the dataset. With the full datasets, and based on the median MCMC loop time, fitting the BHM-HF model was on average 2 to 3% faster compared to BHM-EI, while the latter was found to be about 3% faster when only pluviometer stations are considered. More detailed information regarding computing time and MCMC convergence for each algorithm can be found in Sec. S2 in the Supplementary Material.

On the basis of the results presented in this section, we will focus primarily on the BHM-HF model in the subsequent Sections 6.3 and 6.4.

6.3 Current vs. future IFD curves

The second focus of this work is the integration of RCM outputs into the model in order to estimate future IFD curves. To this end, the BHM's process layer includes, for each GEV parameter, an intercept and the RCM-derived covariate only, as in Eq. (17).

For this scenario, we use GEV parameter estimates obtained from a high-resolution (2 km) run of the Weather Research and Forecasting (WRF) model that used a GCM simulation run

	β_0 (intercept)	β_1 (WRF covariate)	α	λ
$\tilde{\mu}$	2.76, **2.85**, 2.96	-0.045, **-0.011**, 0.021	0.041, **0.057**, 0.08	13.19, **18.82**, 27.45
$\tilde{\sigma}^{(24)}$	3.03, **3.17**, 3.3	0.048, **0.092**, 0.14	0.031, **0.043**, 0.063	27.12, **38.74**, 56.93
ξ	0.077, **0.09**, 0.1	-0.023, **-0.011**, 0.00061	0.007, **0.0086**, 0.01	0.07, **0.67**, 2.16
θ	-3.22, **-2.62**, -1.99	-0.027, **0.059**, 0.15	0.4, **0.63**, 1.01	19.94, **40.46**, 70.15
η	0.58, **0.78**, 1.01	-0.0063, **0.0051**, 0.017	0.032, **0.052**, 0.089	36.71, **60.03**, 99.66

Table 2: Estimated parameters for the WRF model obtained with BHM-HF. Posterior means are given in boldface, with the lower and upper limits of the 95% credible intervals given on either side.

under the SRES A2 emissions scenario as boundary conditions (Evans and McCabe, 2010, 2013; Evans *et al.*, 2014). This RCM output is available only within a subset of the Greater Sydney region, with the top plot in Fig. 1 showing the extents of the WRF grid. A total of 469 stations are used within this extent: $S = 234$ pluviometer and $U = 235$ daily stations. Each station contains at least 8 years of precipitation maxima during the period 1990–2009, which corresponds to the time span for which the current-climate WRF output is available. Additionally, the RCM covariates from the WRF model are also available for the time period 2040–2059.

The BHM-HF model is first fitted to the (daily and pluviometer) precipitation data, using the current-climate WRF outputs as covariates. IFD curves at gauged locations, as well as inferred levels at ungauged locations, can then be computed for this baseline period. To assess how the precipitation extremes might change in the future, the future-climate IFD curves (at the same gauged or ungauged locations) can then be inferred by temporally extrapolating the BHM results using the corresponding WRF outputs obtained for the future climate.

Table 2 provides the posterior means of the main model parameters with their associated 95% credible interval. Of most interest with the current model are the regression coefficients for the WRF covariates (β_1). The results in the second column of Table 2 show that the credible intervals for the marginal posterior distributions of $\tilde{\mu}$, ξ, θ and η each contain zero: for these variables, the BHM is here unable to identify a strong relationship between the fitted GEV parameters at the stations and those provided by the WRF outputs. However, the contribution of the coefficient for $\tilde{\sigma}^{(24)}$ can be seen as being different from zero as its credible interval does not include zero. This allows for the investigation of extreme precipitation for the future climate (2040–2059) when making use of the corresponding WRF outputs.

Fig. 2 shows an example of current and future IFD curves (100-year return period) generated at a selected pluviometer station on the New South Wales coastline (coordinates: 33.28°S, 151.40°E). This plot highlights the difference between extreme precipitation characteristics for the two climatic periods of interest, on the basis of the WRF climate model projections. The uncertainty bands associated with each curve (shaded regions in Fig. 2) provide further information regarding the variability associated with this difference.

Future and current-climate return levels (as shown in Fig. 2) can be calculated for all stations across the region of interest. The differences between these future and present IFD curves can then be computed at each such location, and summarised as shown in Fig. 3. At each duration, these plots show the distribution (box plots) of the difference between IFD curves (left) as well as the distribution of the difference between the width of the uncertainty bands (right). These results, here calculated for a 100-year return period, indicate that IFD levels across all durations will on average be about 7% higher in the future climate compared to the present-time return levels, on the basis of the WRF outputs. However, as can be expected from the temporal extrapolation, the uncertainty around the projected curves also increases, by about

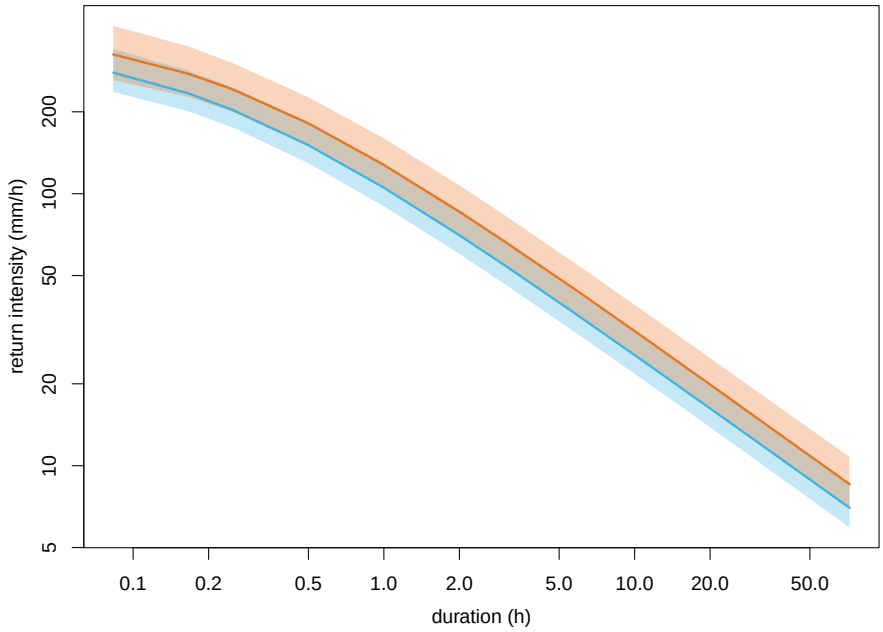

Figure 2: Future (orange) vs. current (blue) IFD curves for a selected pluviometer station in the SYD region (coordinates: 33.28°S, 151.40°E), for a 100-year return period. The shaded regions show the pointwise 95% credible intervals for each curve.

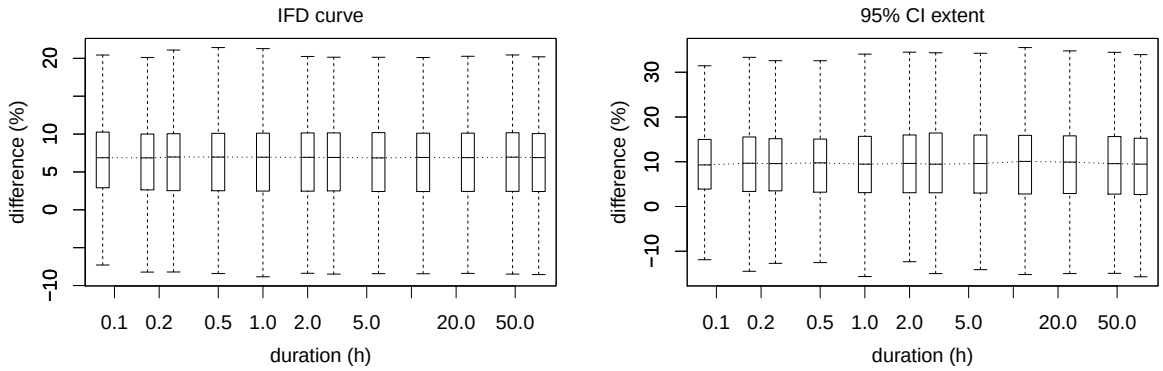

Figure 3: Percentage difference between future and present IFD curves (left) over all stations, and difference between the extent of their respective credible intervals (right), for a 100-year return period.

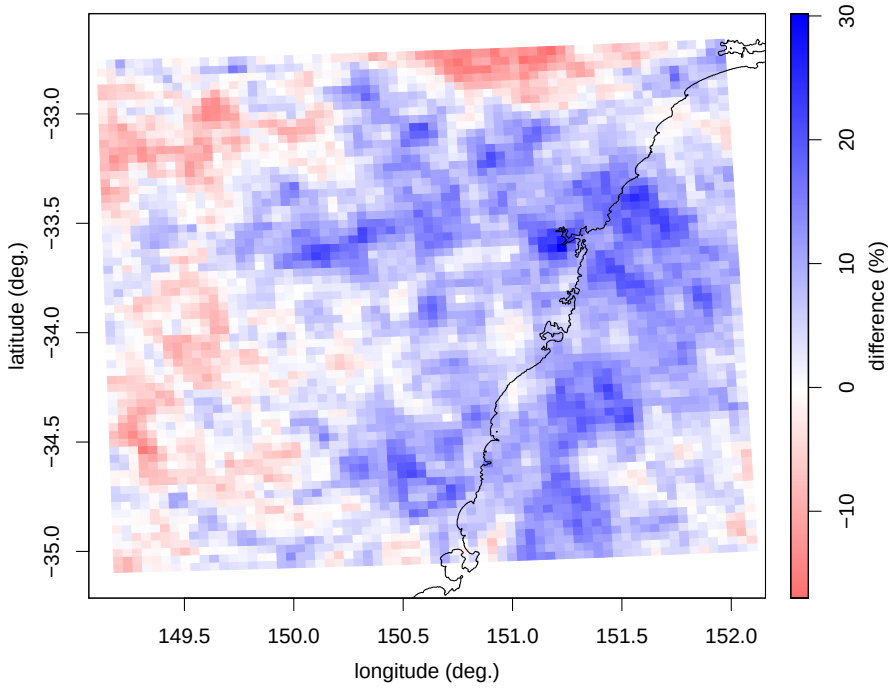

Figure 4: Spatial map of difference between future and current return levels over the study region, for a 100-year return period and a 12h duration. The precipitation extremes become larger in the blue regions, while the red areas essentially become "drier" in the future climate.

10% on average. Similar results (not shown) are obtained with different return periods and for ungauged locations. The BHM-EI approach provides results that are also virtually identical to those presented here.

This type of information (difference in IFD levels) can also be calculated for each grid point where future-climate WRF covariates are available, leading to a spatial map as that shown in Fig. 4. This kind of result highlights the spatial patterns of change in extreme rainfall, here showing that increases in IFD levels are predominantly located along the coast in the study region, while scattered areas of decreased IFD levels can be seen to the west and north. Similar results (not shown) are obtained for other durations, different return periods, and when using BHM-EI.

6.4 Integration of daily and pluviometer stations

One aspect of interest in this work is to consider the effects of jointly considering the daily and pluviometer datasets. To this end, the BHM-HF model is first fitted to the pluviometer data only (310 stations) in the SYD dataset of Sec. 6.2 (Fig. 1, top), following which the IFD curves are inferred at the 562 locations corresponding to the daily stations. These results are then compared to the IFD curves obtained at the same 562 locations when the model is fitted to the whole dataset of pluviometer and daily stations (872 stations in total). The top row in Fig. 5 presents the results as summaries over all 562 IFD curves. It shows that on average, joint modelling of the daily and pluviometer data increases the IFD levels by 10 – 15% where the daily stations are located, while reducing the uncertainty (narrower credible intervals), especially around the 24 h duration where the daily measurements contribute most. Similar trends can also be observed for different return periods, in the SEQ region, and with the BHM-EI algorithm (though with a more pronounced increase in IFD levels overall in comparison with BHM-HF).

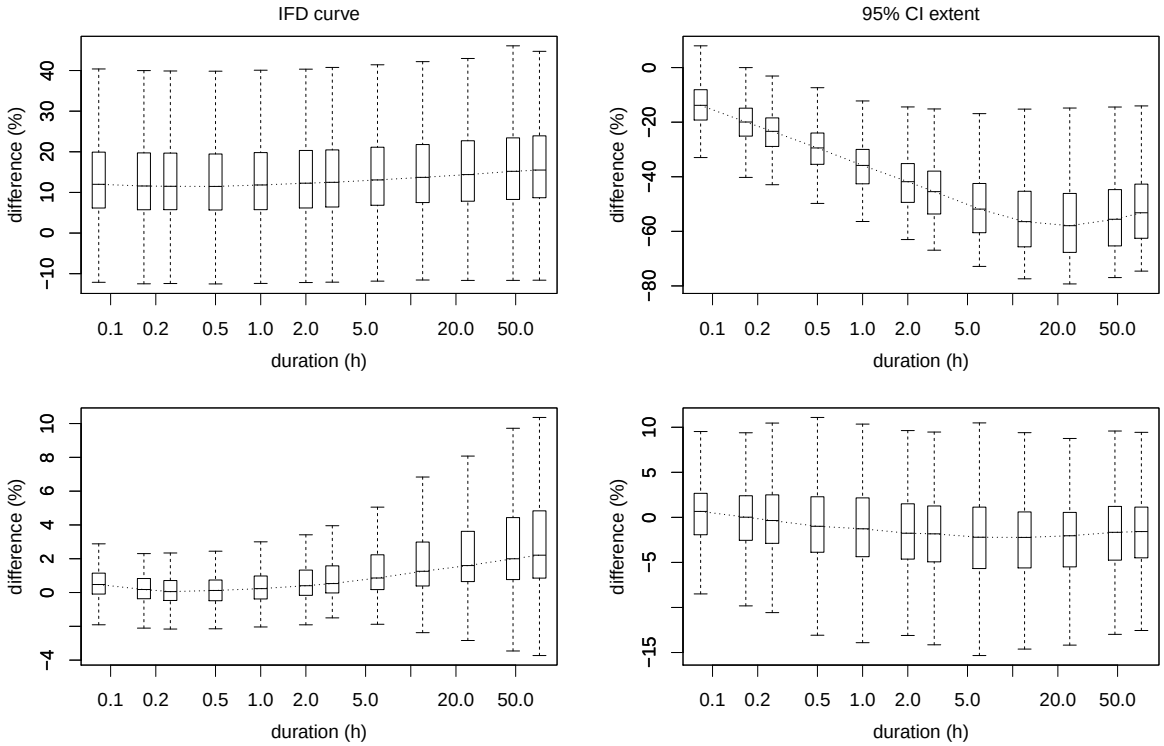

Figure 5: Summary of differences between IFD curves (10-year return period) when daily and pluviometer stations are modelled jointly, compared to IFD levels calculated on the basis of pluviometer data alone: percentage difference between IFD levels (left), and difference in extents of the 95% credible intervals (right). Results calculated at the daily locations (562 IFD curves) are shown in the top row, and at the pluviometer stations (310 IFD curves) in the bottom row.

Note that these results (return levels at the daily stations) are indicative of the largest changes across the study regions. The differences in IFD curves at ungauged locations (not shown) are similar though typically more moderate. Also, the bottom row in Fig. 5 shows the differences in the IFD levels fitted at the pluviometer stations (summary over 310 IFD curves). The effects of the daily stations are here much smaller (between 0 and 2% on average), with the largest differences occurring at longer durations (heaviest contribution from the daily data).

Here again, spatial extrapolation through Bayesian inference at ungauged locations allows for an investigation of the spatial patterns of change associated with the inclusion of daily stations. Fig. 6 presents such results over a grid of points (selected within the hull of stations), indicating that the largest increases in IFD levels are located mostly in the north-western edge of the study region, where the spatial density of daily stations is much larger than that of the pluviometer stations (see Fig. 1, top). Further research would be required to fully understand the (physical or methodological) processes giving rise to these specific change patterns across the study area. The spatial variability of these changes (Fig. 6), as well as the similarity of the trends for both BHM-HF and BHM-EI, would suggest that this is likely a genuine aspect of the data. The meaning of this result should also be considered in the context of each individual IFD curve: further results (not provided here) show that for a large majority of stations and ungauged locations, the IFD curves based on the joint datasets sit within the range of variability (credible interval) of the corresponding return levels based on pluviometer data only. Nevertheless, these results highlight the importance of integrating all the available precipitation datasets into the model, rather than relying on a single source of information.

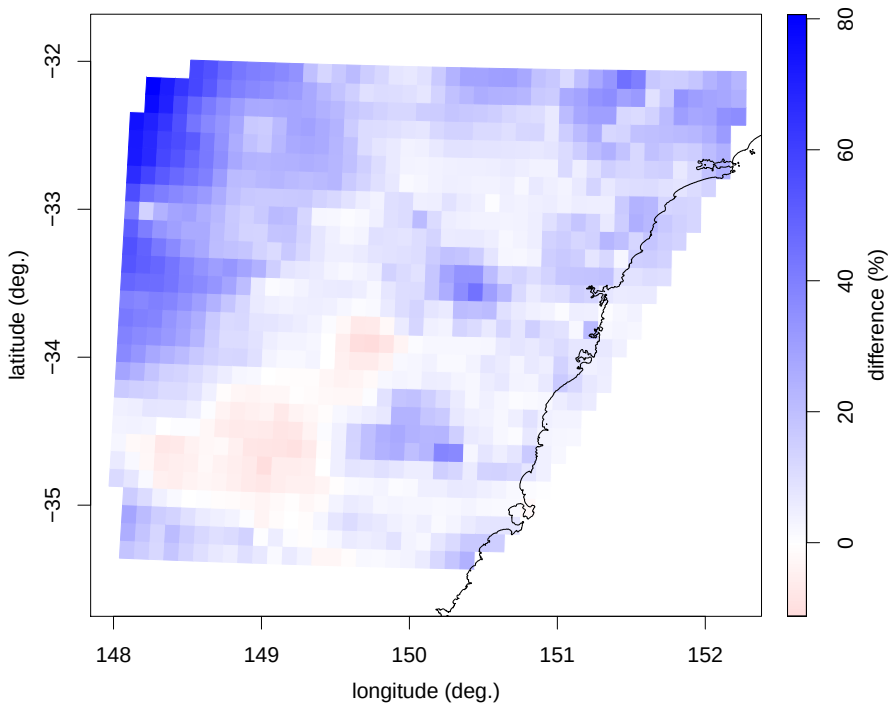

Figure 6: Spatial map of difference between return levels when jointly modelling the pluviometer and daily data (compared to the pluviometer data alone), calculated for a 10-year return period and 12 h duration.

7 Conclusion

This paper presented a flexible and unified framework for the spatial modelling of precipitation extremes over increasing durations. We proposed and assessed two different methods to integrate different sources of rainfall data from pluviometer and daily records. The two approaches were found to yield similar results when applied to two different large-scale study sites in Australia. Model selection and validation tests did not identify one algorithm as being largely superior, although the extremal index approach exhibits a slightly lower predictive accuracy, and tends to produce return levels that are somewhat higher than the model based on the Hershfield factor.

Due to the assumption of conditional independence of the rainfall maxima, the present formulation cannot be used to simulate realistic surfaces of extreme rainfall (Davison *et al.*, 2012; Sang and Gelfand, 2010). Nevertheless, as Cooley and Sain (2012) point out in their discussion of Davison *et al.* (2012), the Bayesian approach used here is sufficient when only the marginal posterior distributions of the GEV parameters are required, as is the case when calculating return levels or IFD curves. Max-stable processes represent an alternative modelling approach that circumvents the restrictions imposed by the conditional independence assumption (Davison *et al.*, 2012; Stephenson *et al.*, 2016).

A key advantage of the Bayesian approach results from the use of a latent spatial model for the GEV parameters. The data are essentially pooled across space (borrowing strength from neighbouring stations) when estimating marginal parameters, leading to more precise estimates of GEV parameters. In addition, the spatial model also allows for a straightforward estimation of parameters at locations other than those at which direct rainfall measurements are available. A similar improvement in estimation is also achieved by using the parametric model of Eq. (2), which combines rainfall data over increasing durations.

Another main advantage of the BHM implementation is that measures of uncertainty can be easily calculated as an intrinsic outcome of the modelling framework. Error characterisation is typically a difficult task with more traditional methods based on a multi-stage approach (e.g. RFA for the calculation of IFD curves; bias-correction or delta methods for the assessment of climate change). Within a BHM, uncertainty at each level of the hierarchy essentially propagates through the model to provide credible intervals on the modelled parameters and related quantities (e.g. return levels and IFD curves).

The integration of various covariates within the proposed approach is straightforward. This includes the use of RCM-based information, with future RCM projections then enabling an investigation of future characteristics of extreme rainfall. Our approach to modelling future-climate IFD curves essentially treats the RCM-based GEV parameters as explanatory covariates integrated into the model at the process level. An alternative approach would be to integrate the RCM's precipitation maxima directly as an additional data layer driven by the underlying GEV parameters of the latent model. Such an approach will be the object of further investigations as part of this research. The proposed framework also lays the groundwork for a more generic approach involving an ensemble (consensus) of climate models, rather than the output from a single RCM as used in this work.

Acknowledgement

We gratefully acknowledge financial support from the Australian Government through Geoscience Australia, and the substantial in-kind support provided by the members of Engineers Australia. Also, we would like to thank the Editor-in-Chief, an anonymous Associate Editor and one anonymous reviewer for their positive feedback and constructive comments on this manuscript.

References

Atyeo J, Walshaw D, 2012. A region-based hierarchical model for extreme rainfall over the UK, incorporating spatial dependence and temporal trend. *Environmetrics* **23**(6): 509–521.

Banerjee S, Carlin B, Gelfand A, 2003. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis.

Bracken C, Rajagopalan B, Cheng L, Kleiber W, Gangopadhyay S, 2016. Spatial Bayesian hierarchical modeling of precipitation extremes over a large domain. *ArXiv e-prints* `arXiv:1512.08560v2 [stat.ME]`.

Brooks SP, Roberts GO, 1998. Convergence assessment techniques for Markov chain Monte Carlo. *Statistics and Computing* **8**: 319–335.

Cheng L, AghaKouchak A, 2014. Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific Reports* **4**(7093).

Coles S, 2001. *An Introduction to Statistical Modeling of Extreme Values*. Lecture Notes in Control and Information Sciences, Springer.

Cooley D, Nychka D, Naveau P, 2007. Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association* **102**(479): 824–840.

Cooley D, Sain SR, 2012. Discussion of "Statistical Modeling of Spatial Extremes" by A. C. Davison, S. A. Padoan and M. Ribatet. *Statistical Science* **27**(2): 187–188.

Davison A, Padoan S, Ribatet M, 2012. Statistical modelling of spatial extremes. *Statistical Science* **27**(2): 161–186.

Dyrrdal AV, Lenkoski A, Thorarinsdottir TL, Stordal F, 2015. Bayesian hierarchical modeling of extreme hourly precipitation in Norway. *Environmetrics* **26**(2): 89–106.

Evans J, McCabe M, 2013. Effect of model resolution on a regional climate model simulation over southeast Australia. *Climate Research* **56**(2): 131–145.

Evans JP, Ji F, Lee C, Smith P, Argüeso D, Fita L, 2014. Design of a regional climate modelling projection ensemble experiment—NARCliM. *Geoscientific Model Development* **7**(2): 621–629.

Evans JP, McCabe MF, 2010. Regional climate simulation over Australia's Murray-Darling basin: A multitemporal assessment. *Journal of Geophysical Research: Atmospheres* **115**(D14).

Garcia-Bartual R, Schneider M, 2001. Estimating maximum expected short-duration rainfall intensities from extreme convective storms. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere* **26**(9): 675–681.

Geirsson ÓP, Hrafnkelsson B, Simpson D, 2015. Computationally efficient spatial modeling of annual maximum 24-h precipitation on a fine grid. *Environmetrics* **26**(5): 339–353.

Ghosh S, Mallick BK, 2011. A hierarchical Bayesian spatio-temporal model for extreme precipitation events. *Environmetrics* **22**(2): 192–204.

Green J, Johnson F, Xuereb K, The C, Moore G, 2012. Revised intensity-frequency-duration (IFD) design rainfall estimates for Australia - An overview. In *Hydrology and Water Resources Symposium*, Sydney, Australia.

Hershfield D, 1961. Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years. Weather Bureau Technical Paper 40, U.S. Department of Commerce, Washington, D.C.

Hershfield D, Wilson W, 1958. Generalizing of rainfall intensity-frequency data. *IUGG/IAHS Publication* **43**: 499–506.

Hosking J, Wallis J, 2005. *Regional Frequency Analysis: An Approach Based on L-Moments.* Cambridge University Press, New York.

IPCC, 2013. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Koutsoyiannis D, Kozonis D, Manetas A, 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. *Journal of Hydrology* **206**(1): 118–135.

Lehmann E, Phatak A, Soltyk S, Chia J, Lau R, Palmer M, 2013. Bayesian hierarchical modelling of rainfall extremes. In *International Congress on Modelling and Simulation*, Adelaide, Australia, 2806–2812.

Muller A, Bacro JN, Lang M, 2008. Bayesian comparison of different rainfall depth-duration-frequency relationships. *Stochastic Environmental Research and Risk Assessment* **22**(1): 33–46.

Nadarajah S, Anderson CW, Tawn JA, 1998. Ordered multivariate extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **60**(2): 473–496.

Peck A, Prodanovic P, Simonovic SPP, 2012. Rainfall intensity duration frequency curves under climate change: City of London, Ontario, Canada. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* **37**(3): 177–189.

Pilgrim D, 1997. *Australian Rainfall and Runoff: a guide to flood estimation*. Institution of Engineers, Australia, Barton, ACT, Australia.

Robert C, Casella G, 2010. *Introducing Monte Carlo Methods with R*. Springer.

Robinson M, Tawn J, 2000. Extremal analysis of processes sampled at different frequencies. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **62**(1): 117–135.

Sang H, Gelfand AE, 2010. Continuous spatial process models for spatial extreme values. *Journal of Agricultural, Biological, and Environmental Statistics* **15**(1): 49–65.

Srivastav R, Schardong A, Simonovic S, 2014. Equidistance quantile matching method for updating IDF curves under climate change. *Water Resources Management* **28**(9): 2539–2562.

Stephenson AG, 2016. Bayesian inference for extreme value modelling. In Dey D, Yan J (eds.), *Extreme value modeling and risk analysis: methods and applications*, Chapman & Hall/CRC, 257–280.

Stephenson AG, Lehmann EA, Phatak A, 2016. A max-stable process model for rainfall extremes at different accumulation durations. Submitted to Weather and Climate Extremes (under revision).

Van de Vyver H, 2012. Spatial regression models for extreme precipitation in Belgium. *Water Resources Research* **48**(9).

Van de Vyver H, 2015. Bayesian estimation of rainfall intensity-duration-frequency relationships. *Journal of Hydrology* **529**(3): 1451–1463.

van Montfort MA, 1997. Concomitants of the Hershfield factor. *Journal of Hydrology* **194**: 357–365.

Wang Y, So M, 2016. A Bayesian hierarchical model for spatial extremes with multiple durations. *Computational Statistics and Data Analysis* **95**: 39–56.

Yilmaz AG, Hossain I, Perera BJC, 2014. Effect of climate change and variability on extreme rainfall intensity-frequency-duration relationships: a case study of Melbourne. *Hydrology and Earth System Sciences* **18**(10): 4065–4076.

Supporting Information

Spatial Modelling Framework for the Characterisation of Rainfall Extremes at Different Durations and under Climate Change

Eric A. Lehmann[1], Aloke Phatak[2], Alec Stephenson[3], Rex Lau[4]

[1]CSIRO Data61, Acton ACT, Australia
[2]Dept. of Mathematics and Statistics, Curtin University, Bentley WA, Australia
[3]CSIRO Data61, Clayton South VIC, Australia
[4]CSIRO Data61, Floreat WA, Australia

This supplementary document provides supporting material for the above paper, detailing some technical aspects of the proposed methodology. The various equations and sections referenced herein can be found in the main article. The sections in this document are not meant to logically follow one another, with the material organised so that the characteristics of the datasets and proposed methods are presented first (Sections S1 – S3), while more mathematical derivations are given at the end (Sections S4 and S5).

S1 Bayesian inference and prior distributions

Based on the notation used in Sec. 3, and defining the set $\boldsymbol{\Omega}$ of parameters as:

$$\boldsymbol{\Omega} = \bigcup_{\chi \in \{\tilde{\mu},\sigma,\xi,\theta,\eta\}} \{\alpha_\chi, \lambda_\chi, \boldsymbol{\beta}_\chi\} \tag{S1}$$

the posterior distribution of the model parameters is determined as follows:

$$p(\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\Omega} | \mathbf{Y}) \propto p(\mathbf{Y}, \tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\Omega}) \tag{S2}$$

$$= p(\mathbf{Y} | \tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}) \cdot p(\boldsymbol{\Omega}) \cdot \prod_{\chi \in \{\tilde{\mu},\sigma,\xi,\theta,\eta\}} p(\boldsymbol{\chi} | \boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi) \tag{S3}$$

where $p(\boldsymbol{\Omega})$ represents the prior density function:

$$p(\boldsymbol{\Omega}) = \prod_{\chi \in \{\tilde{\mu},\sigma,\xi,\theta,\eta\}} p(\boldsymbol{\beta}_\chi)\ p(\alpha_\chi)\ p(\lambda_\chi). \tag{S4}$$

Following Davison *et al.* (2012), we use conjugate Gamma, inverse Gamma and multivariate-normal priors, respectively:

$$\lambda_\chi \sim \text{Gamma}(\kappa_{\lambda_\chi}, \gamma_{\lambda_\chi}) \tag{S5}$$

$$\alpha_\chi \sim \text{InvGamma}(\kappa_{\alpha_\chi}, \gamma_{\alpha_\chi}) \tag{S6}$$

$$\boldsymbol{\beta}_\chi \sim \text{MVN}(\boldsymbol{\mu}_{\beta_\chi}, \boldsymbol{\Sigma}_{\beta_\chi}) \tag{S7}$$

where $\kappa_{(\cdot)}$ and $\gamma_{(\cdot)}$ represent the shape and scale hyper-parameters of the respective distributions, and $\boldsymbol{\mu}_{\beta_\chi}$ and $\boldsymbol{\Sigma}_{\beta_\chi}$ are the prior mean and variance matrix of $\boldsymbol{\beta}_\chi$, respectively.

Inferences from the posterior distribution in Eq. (S3) are obtained using Markov chain Monte Carlo (MCMC) simulation, with implementation details provided in Sec. S5. Selection of the hyper-parameters in our implementation was performed as described in Davison *et al.* (2012). Priors for the regression coefficients $\boldsymbol{\beta}_\chi$ were set to have very large variances. Suitable prior densities were chosen for α_χ and λ_χ on the basis of exploratory analyses of fitted marginal distributions: the priors were set to have means equal to the average marginal maximum likelihood estimates (MLEs), with much larger variances.

S2 Implementation and computational aspects

This section provides specific implementation details for the BHM-HF and BHM-EI methods, as well as their computational requirements when applied to the SYD and SEQ datasets. The following results are based on the models described in Sections 3 and 4 of the main manuscript, with elevation and geographical location used as covariates. The results are based on 150,000 iterations of the MCMC algorithms for each model parameter, with the first 10,000 iterations discarded (burn-in), and thinned by a factor 40. Step sizes for proposal samples within the Metropolis–Hastings steps were set so as to ensure an acceptance rate of about 45%.

For each dataset, MCMC sampling is used to fit a total of P model parameters:

$$P = S \cdot 5 + U \cdot 3 + \sum_{\chi \in \{\tilde{\mu}, \tilde{\sigma}^{(24)}, \xi, \theta, \eta\}} (N_\chi + 1) + 10 \tag{S8}$$

where the first term on the right-hand side represents the number of parameters for the pluviometer stations, the second term is that of the daily stations, the third term corresponds to the regression coefficients ($\boldsymbol{\beta}$) for N_χ covariates and an intercept term, and the last term relates to the spatial model parameters (α and λ). This results in a total of $P_{\text{SYD}} = 3266$ model parameters for the SYD dataset, and $P_{\text{SEQ}} = 4554$ for SEQ.

S2.1 Chain diagnostics

Table S1 provides the lag for which the autocorrelation in the MCMC chains drops below a negligible level. To summarise the results obtained for a large number of model parameters (P), we present averaged values for three different categories, namely GEV parameters, regression coefficients and spatial process parameters. Further results (not presented here) show that the corresponding integrated autocorrelation times averaged over these categories fall within the range between 1.15 and 2.28. These results indicate that an acceptably low autocorrelation is achieved for the chains of all parameters on average.

Convergence results for the same categories of model variables are provided in Table S2. It presents the percentage of MCMC chains having passed the Geweke (Geweke, 1992) and Heidelberger–Welch (Heidelberger and Welch, 1983) diagnostic tests of convergence. These results show no evidence of systematic convergence issues with the simulated algorithms. Results from the Raftery–Lewis diagnostic (Raftery and Lewis, 1992b), not included here, show that the average dependence factor for all parameter categories falls within the range between 1.02 and 1.5, thereby pointing to satisfactory chain lengths and mixing properties; typically, dependence factors larger than 5 would point to a potential convergence failure (Raftery and Lewis, 1992a).

	SYD		SEQ	
	BHM-HF	BHM-EI	BHM-HF	BHM-EI
GEV parameters – pluviometer	2.8 (1550)	2.6 (1550)	2.4 (1200)	2.4 (1200)
GEV parameters – daily	1.6 (1686)	2.0 (1686)	1.7 (3324)	2.3 (3324)
Regression coefficients (20)	1.6	2.6	1.7	3.0
Spatial model parameters (10)	4.7	3.2	2.7	4.2

Table S1: Average lag for which the MCMC chain autocorrelation becomes negligible. Results are averaged for three different categories of parameters: GEV parameters $\tilde{\mu}$, $\tilde{\sigma}^{(24)}$, ξ, θ and η (further split between daily and pluviometer stations), regression coefficients $\boldsymbol{\beta}$, and spatial process parameters α and λ. Numbers in parentheses indicate how many parameters are used in the averaging.

	SYD				SEQ			
	BHM-HF		BHM-EI		BHM-HF		BHM-EI	
	GD	HD	GD	HD	GD	HD	GD	HD
GEV parameters – pluviometer	94.3	99.2	93.2	99.4	94.1	99.3	95.3	99.6
GEV parameters – daily	94.9	99.4	94.8	99.2	94.1	99.4	94.5	99.2
Regression coefficients	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Spatial model parameters	100.0	100.0	90.0	100.0	100.0	90.0	100.0	100.0

Table S2: Proportion of model variables whose MCMC chain passed the Geweke (GD) and Heidelberger–Welch (HD) diagnostic test. All values are given in % as averages over the respective parameter categories (the number of parameters in each category is provided in Table S1).

S2.2 Computational requirements

The BHM-HF and BHM-EI algorithms were implemented in R (R Core Team, 2015) and optimised for computational efficiency. Parallelisation of the main MCMC sampling loop was considered (sampling of the $\boldsymbol{\beta}_{(\cdot)}$, $\alpha_{(\cdot)}$ and $\lambda_{(\cdot)}$ coefficients can be parallelised) but did not substantially reduce the computational efforts due to large overheads. An implementation using block sampling of the model variables was also tested but did not lead to substantial improvements in execution times either: the computational advantages gained through block sampling were offset manyfold by much longer convergence times.

The BHM fitting routines were thus executed sequentially on a single processing core (of a Dual Xeon 8-core E5-2650 2GHz compute node) on CSIRO's high-performance computing infrastructure 'Bragg'. Table S3 shows the detail of the resulting computing requirements for each of the algorithms applied to each of the selected regions of interest. Long chain lengths were specifically used in this work to monitor and investigate potential model convergence issues. Without major degradation in model performance, it is expected that these computation times can be reduced considerably by using a combination of fewer MCMC iterations, a smaller thinning factor (MacEachern and Berliner, 1994), and an implementation based on the simulation of multiple parallel chains.

Once the model parameters are fitted, further calculations (such as IFD curves at ungauged locations, future-climate IFD curves, or gridded outputs of return levels) can be parallelised and were executed on a 16-core node of the Bragg computing infrastructure. Such results were typically achieved in a matter of minutes, up to a maximum of one hour (depending on the number of locations involved).

	SYD		SEQ	
	BHM-HF	BHM-EI	BHM-HF	BHM-EI
median loop time (s)	1.752	1.776	2.760	2.878
total CPU time (h)	75.58	82.52	135.20	131.98

Table S3: Average MCMC loop time and total processing time (based on a total of 150,000 iterations) for model fitting of $P_{\text{SYD}} = 3266$, respectively $P_{\text{SEQ}} = 4554$ parameters.

S3 Model evaluation

S3.1 Koutsoyiannis *et al.* relationship

Exploratory data analysis was carried out on our precipitation datasets in order to assess the dependence model across duration provided in Koutsoyiannis *et al.* (1998). Fig. S1 provides an overview of these results for four randomly selected pluviometer stations in the SEQ dataset. The black dots (with error bars) show the MLEs of the GEV parameters obtained for each duration independently. The red lines are based on MLEs of the GEV parameters using the Koutsoyiannis *et al.* parameterisation of Sec. 2.2 to combine all the durations into a single composite likelihood. Both of these results are obtained for each station separately.

The Koutsoyiannis *et al.* model defines the GEV parameters $\tilde{\mu}$ and ξ as constant across duration, while the duration dependence for $\tilde{\sigma}$ is modelled as per Eq. (2). The plots in Fig. S1 indicate that this dependence model can be reasonably assumed to be valid for the considered stations. With similar results obtained for a vast majority of stations in both the SYD and SEQ datasets, the Koutsoyiannis *et al.* relationship used in Sec. 2.2 was thus found to provide an adequate fit for our data.

For insight, the plots in Fig. S1 also include the Bayesian estimates (with 95% credible intervals) of the same parameters obtained with the BHM-HF method. These model-based estimates draw on the information available across durations and neighbouring stations (through the spatial model), and consequently differ from the MLEs in general, most notably by exhibiting substantially reduced levels of uncertainty. Similar results are obtained with BHM-EI.

S3.2 Cumulative distribution plots

Cumulative distribution function (CDF) plots can be used to provide insight into how well the extreme rainfall data fit under a specific model. Plotting the empirical CDF of the data together with the corresponding model-estimated CDF allows the identification of potential discrepancies between the data and the model outputs.

The dataset $\mathbf{Y}$ used in this work is bi-variate, corresponding to the top $r = 2$ maxima in the time series of precipitation data. A CDF display would thus involve three-dimensional plots of cumulative distribution surfaces over the $(y^{(1)}, y^{(2)})$ domain, leading to considerable plotting and interpretation difficulties. Instead, here we present results in terms of the marginal distributions for $y^{(1)}$, i.e. the top maximum only. On the basis of the r-largest density $\text{GEV}_r(\mathbf{y}; \tilde{\mu}, \tilde{\sigma}, \xi)$, it can be shown that the marginal density for $y^{(1)}$ is the univariate GEV density with same parameters, i.e. $\text{GEV}(y^{(1)}; \tilde{\mu}, \tilde{\sigma}, \xi)$.

Fig. S2 presents the BHM-HF results for three pluviometer stations in the SYD dataset, selected to represent a variety of geographical locations and number of available data points (years). The CDFs are evaluated at four different durations, namely 5 min, 1, 12 and 48 h, as a representative subset of the range of durations in the dataset. These plots indicate that the model provides a satisfactory fit of the data. Similar results (not shown here) are achieved for

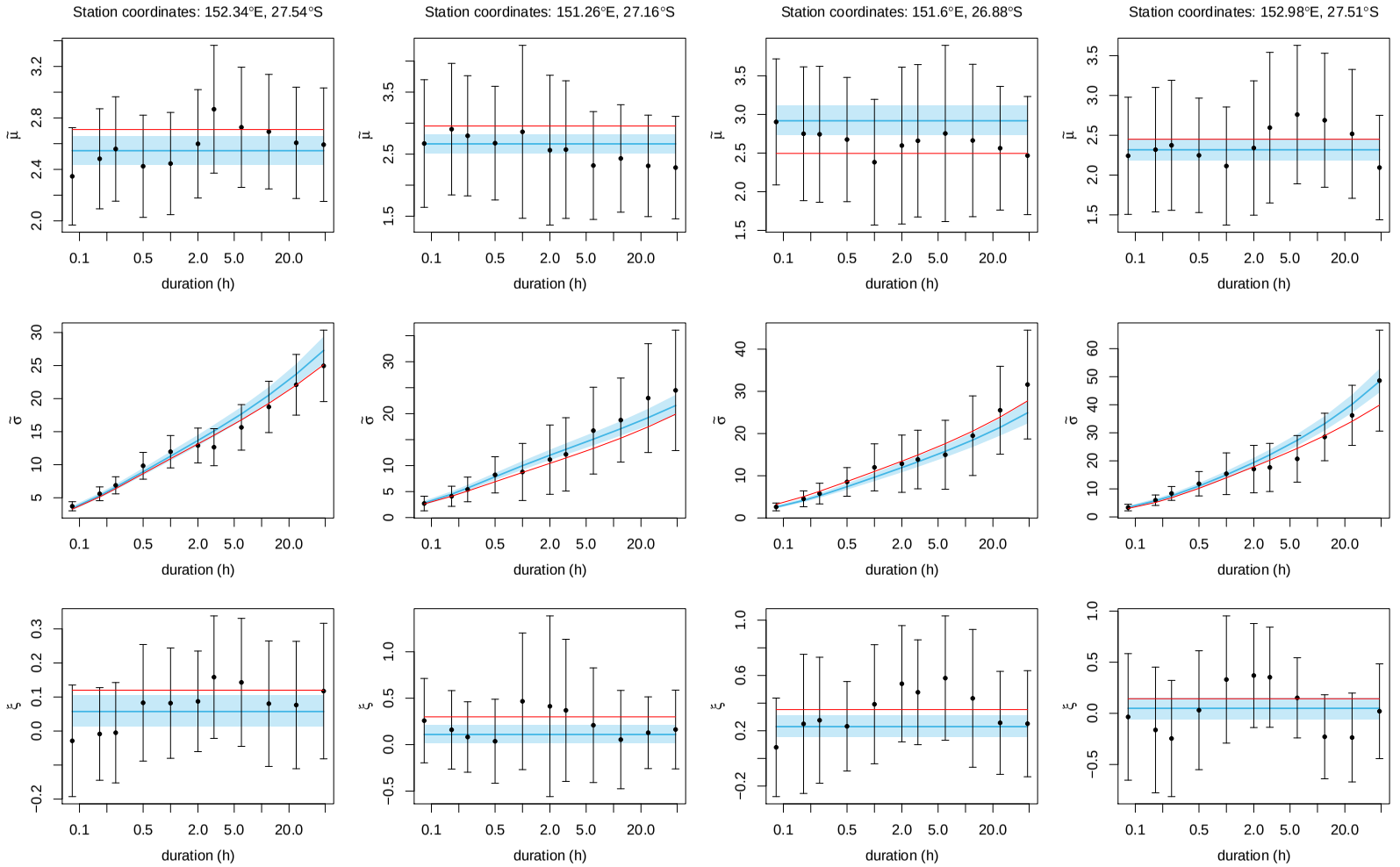

Figure S1: Estimates of GEV parameters $\tilde{\mu}$, $\tilde{\sigma}$ and ξ (rows) for four pluviometer stations (columns) in the SEQ dataset. The black dots with error bars (asymptotic 95% confidence intervals) show the MLEs for each duration separately, while the red lines represent MLEs of the GEV parameters using the Koutsoyiannis *et al.* relationship (each station being considered independently). The blue lines show the model-based estimates (with shaded regions representing the 95% posterior credible intervals) obtained with BHM-HF.

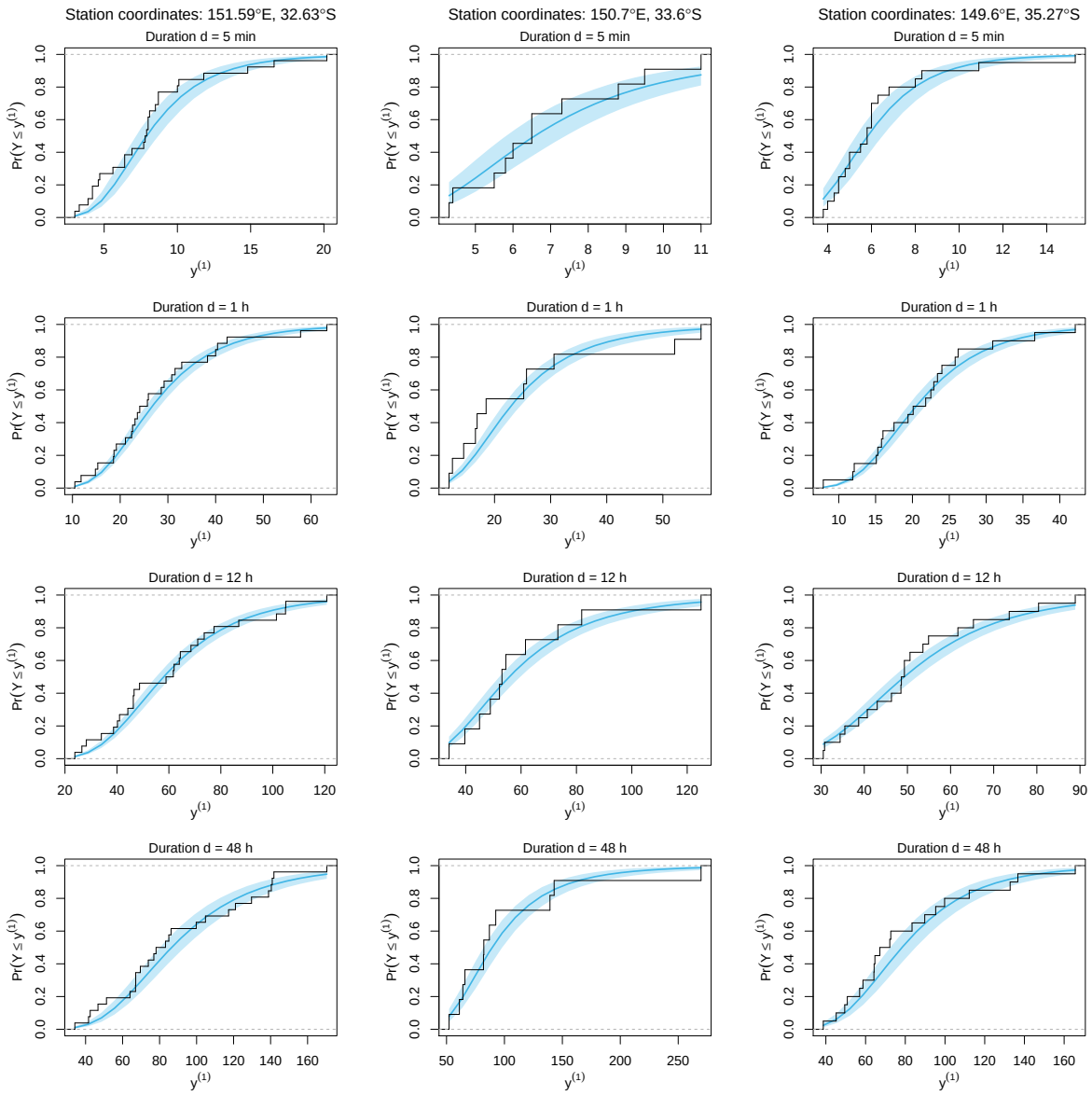

Figure S2: CDF plots for three pluviometer stations (columns) in the SYD dataset and four durations (rows). The black lines show the empirical CDFs for the top maximum data ($y^{(1)}$). The blue lines show the model-based CDFs (with shaded regions representing the 95% posterior credible intervals) obtained with the BHM-HF algorithm.

other durations and other pluviometer stations, as well as for the BHM-EI algorithm and the SEQ dataset.

S3.3 Comparison of posterior densities

Fig. S3 shows plots of the posterior densities (kernel estimates) for each of the main model parameters, for both regions and both approaches (BHM-HF and BHM-EI). Results in this figure suggest that the geographical covariates (easting, northing and elevation in the second, third and fourth columns, respectively) have more inferential power in the SYD region (blue curves), with about 11 of the fitted regression coefficients ($\boldsymbol{\beta}$) being different from zero (i.e. credible intervals exclude zero), compared to about 5 in the SEQ region (red curves).

BHM-HF (solid lines) and BHM-EI (dashed lines) yield similar results for most of the parameters, and particularly for the scale parameter $\tilde{\sigma}^{(24)}$, which represents the main influencing factor in the calculation of IFD curves. Some differences can be seen for estimates related to $\tilde{\mu}$, with the most notable difference occurring with $\lambda_{\tilde{\mu}}$. Here, the BHM-HF results point to a high level of spatial correlation for $\tilde{\mu}$, whereas $\lambda_{\tilde{\mu}}$ becomes virtually negligible under BHM-EI. In contrast, the exact opposite can be seen to occur for λ_{ξ}.

S4 Model selection and validation

Fitting a Bayesian hierarchical model should be followed by an assessment of how well the model fits the data, and of its adequacy in the application for which it was implemented. In this section, we review some aspects of model checking and model comparison which we apply to our BHMs. The following descriptions are mostly based on the review presented in Stephenson (2016, Supplement section) and the references therein, adapted specifically to the Bayesian models proposed in this work.

S4.1 Posterior predictive checking

Posterior predictive checking is a diagnostic that can be used to determine whether the observed data $\mathbf{Y}$ looks plausible under the posterior predictive distribution $p(\check{\mathbf{Y}}|\mathbf{Y})$ (Stephenson, 2016; Gelman *et al.*, 2015). If the model is appropriate, then the simulated data samples $\check{\mathbf{Y}}$ are expected to be similar to the observed maxima $\mathbf{Y}$.

Assume that the MCMC sampling procedure produces K draws (MCMC chains) of the model parameters:

$$\left\{\tilde{\boldsymbol{\mu}}_{+}^{(k)}, \tilde{\boldsymbol{\sigma}}_{+}^{(24)(k)}, \boldsymbol{\xi}_{+}^{(k)}, \boldsymbol{\theta}^{(k)}, \boldsymbol{\eta}^{(k)}\right\}, \qquad k = 1, \ldots, K \tag{S9}$$

from the respective marginal posterior distributions. Simulated data samples $\check{\mathbf{y}}^{(k)}$ from the posterior predictive distribution can then simply be obtained by drawing from

$$p\big(\check{\mathbf{Y}}|\tilde{\boldsymbol{\mu}}_{+}^{(k)}, \tilde{\boldsymbol{\sigma}}_{+}^{(24)(k)}, \boldsymbol{\xi}_{+}^{(k)}, \boldsymbol{\theta}^{(k)}, \boldsymbol{\eta}^{(k)}\big) \tag{S10}$$

for each $k = 1, \ldots, K$.

For pluviometer station $s \in \{1, \ldots, S\}$, year $t \in \{1, \ldots, T_s\}$, and accumulation duration $d_{s,i}$, $i \in \{1, \ldots, D_s\}$, posterior predictive samples are thus generated according to:

$$\check{\mathbf{y}}_{s,t,i}^{(k)} \sim \mathrm{GEV}_r^{*}\left(\tilde{\mu}_s^{(k)}, \tilde{\sigma}_s^{(24)(k)}, \xi_s^{(k)}, \theta_s^{(k)}, \eta_s^{(k)}\right). \tag{S11}$$

For daily stations $u \in \{S+1, \ldots, S+U\}$, generating the posterior predictive samples depends on the considered model type. For BHM-HF, we have:

$$H \cdot \check{\mathbf{y}}_{u,t}^{(\mathrm{D})(k)} \sim \mathrm{GEV}_r\left(\tilde{\mu}_u^{(k)}, \tilde{\sigma}_u^{(24)(k)}, \xi_u^{(k)}\right) \tag{S12}$$

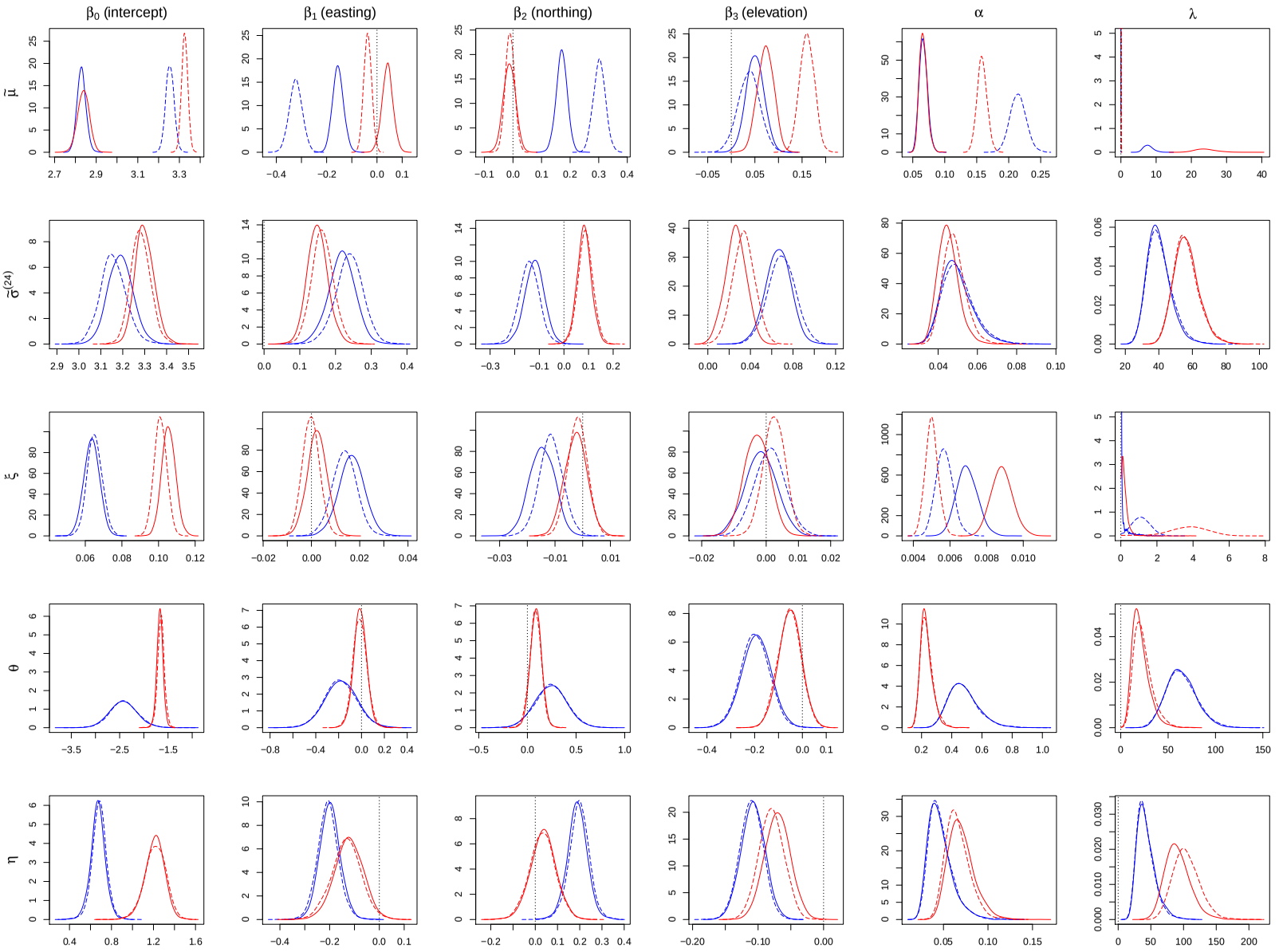

Figure S3: Estimated posterior densities of model parameters for the SYD region (blue) and SEQ region (red), using BHM-HF (solid lines) and BHM-EI (dashed lines). Each column represents a different coefficient for each GEV parameter (one per row).

while the following sampling process is used for BHM-EI:

$$\check{\mathbf{y}}_{u,t}^{(\mathrm{D})(k)} \sim \mathrm{GEV}_r\left(\tilde{\mu}_u^{(\mathrm{D})(k)}, \sigma_u^{(\mathrm{D})(k)}, \xi_u^{(\mathrm{D})(k)}\right). \tag{S13}$$

A comparison of the K posterior predictive samples $\{\check{\mathbf{y}}^{(k)}\}$ with the observed data $\mathbf{Y}$ can be achieved using a specific test statistic $\tau(\cdot)$, and calculating the proportion p_c of simulated samples for which the test statistic is greater than the value calculated for the actual data. A value of p_c close to 0 or 1 indicates that the test statistic corresponds to a feature that is poorly fitted by the model. $\tau(\cdot)$ should thus be defined to reflect aspects of the model that are relevant to the purposes for which the inference is applied, and for extreme value models, Stephenson (2016) suggests defining $\tau(\cdot)$ on the basis of the maximum value. In this work, we thus use:

$$p_c = \frac{1}{r \cdot K \cdot \left(\sum_{s=1}^{S} D_s + U\right)} \times \sum_{j=1}^{r}\sum_{k=1}^{K}\left[\sum_{s=1}^{S}\sum_{i=1}^{D_s}\tau\left(\left[\check{\mathbf{y}}_{s,t,i}^{(k)}\right]_{(j)}, \left[\mathbf{y}_{s,t,i}\right]_{(j)}\right) + \sum_{u=S+1}^{S+U}\tau\left(\left[\check{\mathbf{y}}_{u,t}^{(\mathrm{D})(k)}\right]_{(j)}, \left[\mathbf{y}_{u,t}^{(\mathrm{D})}\right]_{(j)}\right)\right] \tag{S14}$$

with the following test statistic applied to vectors of precipitation maxima for T_s (respectively T_u) years:

$$\tau(\mathbf{x}, \mathbf{y}) = \begin{cases} 1 & \text{if } \max(\mathbf{x}) > \max(\mathbf{y}) \\ 0 & \text{otherwise.} \end{cases} \tag{S15}$$

S4.2 Posterior predictive coverage

An alternative approach to the task of posterior predictive checking is to consider how well the observed data $\mathbf{Y}$ fits under the estimated posterior predictive density. A measure of this fit is to provide the percentage of observed data points contained within the 95% probability interval (highest posterior density) of $p\big(\mathbf{Y}|\hat{\tilde{\boldsymbol{\mu}}}_+, \hat{\tilde{\boldsymbol{\sigma}}}_+^{(24)}, \hat{\boldsymbol{\xi}}_+, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}}\big)$, with $\hat{\boldsymbol{\chi}}$ representing the estimated posterior mean for parameter $\boldsymbol{\chi} \in \{\tilde{\boldsymbol{\mu}}_+, \tilde{\boldsymbol{\sigma}}_+^{(24)}, \boldsymbol{\xi}_+, \boldsymbol{\theta}, \boldsymbol{\eta}\}$:

$$\hat{\boldsymbol{\chi}} = \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{\chi}^{(k)}. \tag{S16}$$

For pluviometer station $s \in \{1, \ldots, S\}$ at accumulation duration $d_{s,i}$, $i \in \{1, \ldots, D_s\}$, the predictive coverage score is calculated by determining how many of the data points $\mathbf{y}_{s,t,i}$, $t \in \{1, \ldots, T_s\}$, fall within the 95% probability region of $\mathrm{GEV}_r^*\big(\hat{\tilde{\mu}}_s, \hat{\tilde{\sigma}}_s^{(24)}, \hat{\xi}_s, \hat{\theta}_s, \hat{\eta}_s\big)$. For daily station $u \in \{S+1, \ldots, S+U\}$, the coverage of $\mathbf{y}_{u,t}^{(\mathrm{D})}$, $t \in \{1, \ldots, T_u\}$, is checked against $\mathrm{GEV}_r\big(\hat{\tilde{\mu}}_u, \hat{\tilde{\sigma}}_u^{(24)}, \hat{\xi}_u\big)$ for BHM-HF, or $\mathrm{GEV}_r\big(\hat{\tilde{\mu}}_u^{(\mathrm{D})}, \hat{\sigma}_u^{(\mathrm{D})}, \hat{\xi}_u^{(\mathrm{D})}\big)$ for BHM-EI. The contributions from all stations are then pooled to calculate a single overall percentage score for the posterior predictive coverage of observed data for a given model.

S4.3 Predictive accuracy score

A measure of predictive accuracy can be obtained by setting aside Q validation stations (randomly selected), and fitting each model to the rest of the precipitation dataset ("training" data, with $S + U - Q$ stations). In this work, we only select pluviometer stations to provide the validation data: $\check{\mathbf{y}}_{q,t,i}$, $q = 1, \ldots, Q$.

Once the model is fitted to the training dataset, spatial interpolation of the results is carried out through the BHM's process layer to infer the value of the GEV parameters:

$$\left\{\breve{\tilde{\mu}}_q^{(k)}, \breve{\tilde{\sigma}}_q^{(24)(k)}, \breve{\xi}_q^{(k)}, \breve{\theta}_q^{(k)}, \breve{\eta}_q^{(k)}\right\} \tag{S17}$$

at the locations of the validation stations $q = 1, \ldots, Q$, for each MCMC iteration $k = 1, \ldots, K$. Given these inferred parameters, a global measure of predictive accuracy can be obtained by averaging the accuracy "score" achieved for each individual validation data point (Stephenson, 2016):

$$\text{pred. acc.} = \frac{1}{K \cdot \sum_{q=1}^{Q} D_q T_q} \sum_{q=1}^{Q} \sum_{t=1}^{T_q} \sum_{i=1}^{D_q} \sum_{k=1}^{K} \text{GEV}_r^* \left(\breve{\mathbf{y}}_{q,t,i}; \breve{\tilde{\mu}}_q^{(k)}, \breve{\tilde{\sigma}}_q^{(24)(k)}, \breve{\xi}_q^{(k)}, \breve{\theta}_q^{(k)}, \breve{\eta}_q^{(k)}\right). \tag{S18}$$

This result provides a measure of the overall fit of the validation data, and can thus be used for model comparison purposes, with larger values indicating better models.

S4.4 Deviance information criterion (DIC)

Different models can also be compared by means of information criteria, which can be regarded as a simplified approach for the evaluation of predictive accuracy (Stephenson, 2016). One such measure is the deviance information criterion (Spiegelhalter *et al.*, 2002), which is defined as follows for the considered models:

$$\text{DIC} = 2 \cdot \log p\left(\mathbf{Y} \middle| \hat{\tilde{\boldsymbol{\mu}}}_+, \hat{\tilde{\boldsymbol{\sigma}}}_+^{(24)}, \hat{\boldsymbol{\xi}}_+, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}}\right) - \frac{4}{K} \sum_{k=1}^{K} \log p\left(\mathbf{Y} \middle| \tilde{\boldsymbol{\mu}}_+^{(k)}, \tilde{\boldsymbol{\sigma}}_+^{(24)(k)}, \boldsymbol{\xi}_+^{(k)}, \boldsymbol{\theta}^{(k)}, \boldsymbol{\eta}^{(k)}\right). \tag{S19}$$

The composite likelihood function, applicable to both terms in the above equation, is defined as follows:

$$p\left(\mathbf{Y} | \tilde{\boldsymbol{\mu}}_+, \tilde{\boldsymbol{\sigma}}_+^{(24)}, \boldsymbol{\xi}_+, \boldsymbol{\theta}, \boldsymbol{\eta}\right) = \prod_{s=1}^{S} \prod_{t=1}^{T_s} \prod_{i=1}^{D_s} p\left(\left[\mathbf{Y}^{(\text{P})}\right]_{(s,t,i)} \middle| \tilde{\boldsymbol{\mu}}_+, \tilde{\boldsymbol{\sigma}}_+^{(24)}, \boldsymbol{\xi}_+, \boldsymbol{\theta}, \boldsymbol{\eta}\right) \times \prod_{u=S+1}^{S+U} \prod_{t=1}^{T_u} p\left(\left[\mathbf{Y}^{(\text{D})}\right]_{(u,t)} \middle| \tilde{\boldsymbol{\mu}}_+, \tilde{\boldsymbol{\sigma}}_+^{(24)}, \boldsymbol{\xi}_+\right) \tag{S20}$$

where we have

$$p\left(\left[\mathbf{Y}^{(\text{P})}\right]_{(s,t,i)} \middle| \tilde{\boldsymbol{\mu}}_+, \tilde{\boldsymbol{\sigma}}_+^{(24)}, \boldsymbol{\xi}_+, \boldsymbol{\theta}, \boldsymbol{\eta}\right) = \text{GEV}_r^*\left(\mathbf{y}_{s,t,i}; \tilde{\mu}_s, \tilde{\sigma}_s^{(24)}, \xi_s, \theta_s, \eta_s\right) \tag{S21}$$

$$p\left(\left[\mathbf{Y}^{(\text{D})}\right]_{(u,t)} \middle| \tilde{\boldsymbol{\mu}}_+, \tilde{\boldsymbol{\sigma}}_+^{(24)}, \boldsymbol{\xi}_+\right) = \begin{cases} H^r \cdot \text{GEV}_r\left(H \cdot \mathbf{y}_{u,t}^{(\text{D})}; \tilde{\mu}_u, \tilde{\sigma}_u^{(24)}, \xi_u\right) & \text{for BHM-HF} \\ \text{GEV}_r\left(\mathbf{y}_{u,t}^{(\text{D})}; \tilde{\mu}_u^{(\text{D})}, \sigma_u^{(\text{D})}, \xi_u^{(\text{D})}\right) & \text{for BHM-EI.} \end{cases} \tag{S22}$$

The DIC parameter is defined on the deviance scale, with smaller DIC values indicating better models.

S4.5 Widely applicable information criterion (WAIC)

An alternative criterion for model comparison is the widely applicable information criterion (Watanabe, 2010). For the proposed models, this criterion is given by

$$\begin{aligned}\mathrm{WAIC} = &\sum_{s=1}^{S}\sum_{t=1}^{T_s}\sum_{i=1}^{D_s}\left\{2\cdot\log\left[\frac{1}{K}\sum_{k=1}^{K} p\left(\left[\mathbf{Y}^{(\mathrm{P})}\right]_{(s,t,i)}\middle|\tilde{\boldsymbol{\mu}}_+^{(k)},\tilde{\boldsymbol{\sigma}}_+^{(24)(k)},\boldsymbol{\xi}_+^{(k)},\boldsymbol{\theta}^{(k)},\boldsymbol{\eta}^{(k)}\right)\right] - \right.\\ &\left.\frac{4}{K}\sum_{k=1}^{K}\log p\left(\left[\mathbf{Y}^{(\mathrm{P})}\right]_{(s,t,i)}\middle|\tilde{\boldsymbol{\mu}}_+^{(k)},\tilde{\boldsymbol{\sigma}}_+^{(24)(k)},\boldsymbol{\xi}_+^{(k)},\boldsymbol{\theta}^{(k)},\boldsymbol{\eta}^{(k)}\right)\right\} + \\ &\sum_{u=S+1}^{S+U}\sum_{t=1}^{T_u}\left\{2\cdot\log\left[\frac{1}{K}\sum_{k=1}^{K} p\left(\left[\mathbf{Y}^{(\mathrm{D})}\right]_{(u,t)}\middle|\tilde{\boldsymbol{\mu}}_+^{(k)},\tilde{\boldsymbol{\sigma}}_+^{(24)(k)},\boldsymbol{\xi}_+^{(k)}\right)\right] - \right.\\ &\left.\frac{4}{K}\sum_{k=1}^{K}\log p\left(\left[\mathbf{Y}^{(\mathrm{D})}\right]_{(u,t)}\middle|\tilde{\boldsymbol{\mu}}_+^{(k)},\tilde{\boldsymbol{\sigma}}_+^{(24)(k)},\boldsymbol{\xi}_+^{(k)}\right)\right\} \end{aligned} \tag{S23}$$

using the likelihood definitions given in Eqs. (S21) and (S22).

S5 Full conditionals

This section provides an overview of the full conditional distributions derived for MCMC implementation of Eq. (S3). Unless otherwise mentioned, standard Gibbs sampling is used. Otherwise, derivation of the acceptance probability is also provided for those variables requiring a Metropolis–Hastings (MH) step within the Gibbs sampler.

The derivations presented below are based on the model of Sec. 3, which assumes a single dataset of S pluviometer stations. Extension of this framework to account for an additional dataset of daily measurements (as described in Sec. 4), using either the Hershfield factor or extremal index approach, is straightforward.

S5.1 GEV parameters $\tilde{\boldsymbol{\mu}}$ and $\boldsymbol{\xi}$

The process model for $\boldsymbol{\xi}$ is similar to that of $\tilde{\boldsymbol{\mu}}$ (see Eq. (6) in Sec. 3.2); the same derivations can thus be used to determine the respective full conditional distribution. Let us define the generic variable $\chi \in \{\tilde{\mu}, \xi\}$ and the related complementary variable:

$$\bar{\chi} \triangleq \begin{cases} \xi & \text{if } \chi \triangleq \tilde{\mu} \\ \tilde{\mu} & \text{if } \chi \triangleq \xi. \end{cases} \tag{S24}$$

With the spatial process definition given as follows (see Sec. 3.2):

$$\mathcal{P}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi) \sim \mathrm{MVN}\big(\mathbf{0}, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi)\big), \tag{S25}$$

the full conditional distribution for the model parameter $\boldsymbol{\chi}$ results from Eqs. (S3), (5) and (6) as:

$$p(\boldsymbol{\chi}|\mathrm{rest}) \propto p(\mathbf{Y}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}) \cdot p(\boldsymbol{\chi}|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi) \tag{S26}$$

$$= \prod_{s=1}^{S}\prod_{t=1}^{T_s}\prod_{i=1}^{D_s} \mathrm{GEV}_r^*(\mathbf{y}_{s,t,i}; \tilde{\mu}_s, \sigma_s, \xi_s, \theta_s, \eta_s) \cdot \mathrm{MVN}\big(\boldsymbol{\chi}; \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi)\big). \tag{S27}$$

As Eq. (S27) does not represent a standard distribution, a MH step is used to update each component of $\boldsymbol{\chi}$ in turn during the MCMC sampling. At the k-th MCMC iteration, a proposal sample is generated for the n-th element of $\boldsymbol{\chi}$, $n \in \{1, \ldots, S\}$, using a symmetric random walk process:

$$\chi_n^* \sim \mathrm{N}\left(\chi_n^{(k-1)}, \sigma^2_{\mathrm{prop},\chi}\right) \tag{S28}$$

where $\chi_n^{(k-1)}$ represents the n-th element of $\boldsymbol{\chi}$ at the previous MCMC iteration, and with $\sigma_{\mathrm{prop},\chi}$ representing the "step size" of the proposal distribution. From the current state $\boldsymbol{\chi}'$ of the $\boldsymbol{\chi}$ vector:

$$\boldsymbol{\chi}' = \left[\chi_1^{(k)}, \ldots, \chi_{n-1}^{(k)}, \chi_n^{(k-1)}, \chi_{n+1}^{(k-1)}, \ldots, \chi_S^{(k-1)}\right]^T \tag{S29}$$

the proposal $\boldsymbol{\chi}$ vector thus results as:

$$\boldsymbol{\chi}^* = \left[\chi_1^{(k)}, \ldots, \chi_{n-1}^{(k)}, \chi_n^*, \chi_{n+1}^{(k-1)}, \ldots, \chi_S^{(k-1)}\right]^T. \tag{S30}$$

The proposal sample χ_n^* is subsequently accepted or rejected according to:

$$\chi_n^{(k)} = \begin{cases} \chi_n^{(k-1)} & \text{with probability } 1 - P_{\mathrm{acc},\chi} \\ \chi_n^* & \text{with probability } P_{\mathrm{acc},\chi} \end{cases} \tag{S31}$$

where the acceptance probability is defined as:

$$P_{\mathrm{acc},\chi} = \min\left\{1, \frac{p(\mathbf{Y}|\boldsymbol{\chi}^*, \bar{\boldsymbol{\chi}}, \boldsymbol{\sigma}, \boldsymbol{\theta}, \boldsymbol{\eta})}{p(\mathbf{Y}|\boldsymbol{\chi}', \bar{\boldsymbol{\chi}}, \boldsymbol{\sigma}, \boldsymbol{\theta}, \boldsymbol{\eta})} \cdot \frac{p(\boldsymbol{\chi}^*|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi)}{p(\boldsymbol{\chi}'|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi)}\right\} \tag{S32}$$

$$= \min\left\{1, \underbrace{\frac{\prod_s \prod_t \prod_i \mathrm{GEV}_r^*\left(\mathbf{y}_{s,t,i}; [\boldsymbol{\chi}^*]_{(s)}, \bar{\chi}_s, \sigma_s, \theta_s, \eta_s\right)}{\prod_s \prod_t \prod_i \mathrm{GEV}_r^*\left(\mathbf{y}_{s,t,i}; [\boldsymbol{\chi}']_{(s)}, \bar{\chi}_s, \sigma_s, \theta_s, \eta_s\right)}}_{A_0} \cdot \underbrace{\frac{\mathrm{MVN}(\boldsymbol{\chi}^*; \mathbf{X}_\chi \boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi)}{\mathrm{MVN}(\boldsymbol{\chi}'; \mathbf{X}_\chi \boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi)}}_{A_1}\right\}. \tag{S33}$$

Since all but one element of the vectors $\boldsymbol{\chi}^*$ and $\boldsymbol{\chi}'$ are identical, the term A_0 in Eq. (S33) can be simplified through some simple algebra:

$$A_0 = \frac{\prod_{t=1}^{T_s} \prod_{i=1}^{D_s} \mathrm{GEV}_r^*\left(\mathbf{y}_{n,t,i}; \chi_n^*, \bar{\chi}_n, \sigma_n, \theta_n, \eta_n\right)}{\prod_{t=1}^{T_s} \prod_{i=1}^{D_s} \mathrm{GEV}_r^*\left(\mathbf{y}_{n,t,i}; \chi_n^{(k-1)}, \bar{\chi}_n, \sigma_n, \theta_n, \eta_n\right)}. \tag{S34}$$

Due to $\boldsymbol{\Sigma}_\chi$ being symmetric, the second fraction A_1 can also be simplified further (leading to a substantial reduction in computational requirements):

$$A_1 = \exp\left\{\left(\mathbf{X}_\chi \boldsymbol{\beta}_\chi - \boldsymbol{\chi}''\right)^T \cdot \left[\boldsymbol{\Sigma}_\chi^{-1}\right]_{(\cdot,n)} \cdot \left(\chi_n^* - \chi_n^{(k-1)}\right)\right\} \tag{S35}$$

where $[\boldsymbol{\Sigma}]_{(\cdot,n)}$ extracts the n-th column of the matrix $\boldsymbol{\Sigma}$, and with

$$\boldsymbol{\chi}'' = \left[\chi_1^{(k)}, \ldots, \chi_{n-1}^{(k)}, \frac{1}{2} \cdot \left(\chi_n^* + \chi_n^{(k-1)}\right), \chi_{n+1}^{(k-1)}, \ldots, \chi_S^{(k-1)}\right]^T. \tag{S36}$$

S5.2 GEV parameters $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$

Similar definitions of the process model for $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$ in Sec. 3.2 allow for the same full conditionals derivations for both parameters. For $\chi \in \{\sigma, \theta\}$ (with $\bar{\chi}$ defined similarly to Eq. (S24)), the full conditional distribution for $\boldsymbol{\chi}$ results from Eq. (S3) as:

$$p(\boldsymbol{\chi}|\mathrm{rest}) \propto p(\mathbf{Y}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}) \cdot p(\boldsymbol{\chi}|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi). \tag{S37}$$

From Eqs. (6) and (S25), we have:

$$\log(\boldsymbol{\chi}) \sim \text{MVN}\big(\mathbf{X}_\chi \boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi)\big) \tag{S38}$$

from which $p(\boldsymbol{\chi}|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi)$ must be derived.

Lemma S1. Given a collection of random variables $(X_1, \ldots, X_N)$ with joint density function $p_{(X_1,\ldots,X_N)}(\cdot)$, it can be shown that the probability density function of the transformed set of variables $(Z_1, \ldots, Z_N) = (f(X_1), \ldots, f(X_N))$ is given by:

$$p_{(Z_1,\ldots,Z_N)}(z_1, \ldots, z_N) = p_{(X_1,\ldots,X_N)}\left(f^{-1}(z_1), \ldots, f^{-1}(z_N)\right) \cdot |J(z_1, \ldots, z_N)| \tag{S39}$$

where $J(z_1, \ldots, z_N)$ represents the Jacobian of the transformation.

Defining $f(\cdot) \triangleq \exp(\cdot)$ and $z_i \triangleq \chi_i$ in conjunction with Eq. (S39), the desired density can thus be derived from Eq. (S38) as:

$$p(\boldsymbol{\chi}|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi) = \text{MVN}\big(\log(\boldsymbol{\chi}); \mathbf{X}_\chi \boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi\big) \cdot \left(\prod_{s=1}^{S} \chi_s\right)^{-1}. \tag{S40}$$

Hence, Eq. (S37) becomes:

$$p(\boldsymbol{\chi}|\text{rest}) \propto \prod_{s=1}^{S}\prod_{t=1}^{T_s}\prod_{i=1}^{D_s} \text{GEV}_r^*(\mathbf{y}_{s,t,i}; \tilde{\mu}_s, \sigma_s, \xi_s, \theta_s, \eta_s) \cdot \text{MVN}\big(\log(\boldsymbol{\chi}); \mathbf{X}_\chi \boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi\big) \cdot \left(\prod_{s=1}^{S} \chi_s\right)^{-1}. \tag{S41}$$

A MH step is used to sample from the non-standard density in Eq. (S41). So as to improve the MCMC sampling convergence, we generate proposal samples $\chi_n^* > 0$ for the n-th component of $\boldsymbol{\chi}$ using a log-normal distribution:

$$p\left(\chi_n^* | \chi_n^{(k-1)}, \sigma^2_{\text{prop},\chi}\right) = \text{logN}\left(\chi_n^*; \log\left(\chi_n^{(k-1)}\right), \sigma^2_{\text{prop},\chi}\right). \tag{S42}$$

The proposal sample is then accepted with probability $P_{\text{acc},\chi}$ defined as:

$$P_{\text{acc},\chi} = \min\left\{1, \underbrace{\frac{p(\mathbf{Y}|\boldsymbol{\chi}^*, \bar{\boldsymbol{\chi}}, \tilde{\boldsymbol{\mu}}, \boldsymbol{\xi}, \boldsymbol{\eta})}{p(\mathbf{Y}|\boldsymbol{\chi}', \bar{\boldsymbol{\chi}}, \tilde{\boldsymbol{\mu}}, \boldsymbol{\xi}, \boldsymbol{\eta})}}_{B_0} \cdot \underbrace{\frac{p(\boldsymbol{\chi}^*|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi)}{p(\boldsymbol{\chi}'|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi)}}_{B_1} \cdot \underbrace{\frac{p\left(\chi_n^{(k-1)}|\chi_n^*, \sigma^2_{\text{prop},\chi}\right)}{p\left(\chi_n^*|\chi_n^{(k-1)}, \sigma^2_{\text{prop},\chi}\right)}}_{B_2}\right\} \tag{S43}$$

with $\boldsymbol{\chi}'$ and $\boldsymbol{\chi}^*$ defined similarly to Eqs. (S29) and (S30), respectively. As per Eq. (S34), the term B_0 becomes:

$$B_0 = \frac{\prod_{t=1}^{T_s}\prod_{i=1}^{D_s} \text{GEV}_r^*\left(\mathbf{y}_{n,t,i}; \chi_n^*, \bar{\chi}_n, \tilde{\mu}_n, \xi_n, \eta_n\right)}{\prod_{t=1}^{T_s}\prod_{i=1}^{D_s} \text{GEV}_r^*\left(\mathbf{y}_{n,t,i}; \chi_n^{(k-1)}, \bar{\chi}_n, \tilde{\mu}_n, \xi_n, \eta_n\right)}. \tag{S44}$$

Analogous to Eq. (S35), and using Eq. (S40), the term B_1 simplifies to:

$$B_1 = \underbrace{\exp\left\{\left(\mathbf{X}_\chi \boldsymbol{\beta}_\chi - \log(\boldsymbol{\chi}'')\right)^T \cdot \left[\boldsymbol{\Sigma}_\chi^{-1}\right]_{(\cdot,n)} \cdot \log\left(\frac{\chi_n^*}{\chi_n^{(k-1)}}\right)\right\}}_{B_3} \cdot \frac{\chi_n^{(k-1)}}{\chi_n^*} \tag{S45}$$

with

$$\boldsymbol{\chi}'' = \left[\chi_1^{(k)}, \ldots, \chi_{n-1}^{(k)}, \sqrt{\chi_n^* \cdot \chi_n^{(k-1)}}, \chi_{n+1}^{(k-1)}, \ldots, \chi_S^{(k-1)}\right]^T. \tag{S46}$$

The term B_2 in Eq. (S43) is introduced to compensate for the use of a non-symmetric proposal distribution. Based on Eq. (S42) and following some algebraic manipulations, this term can be shown to simplify to:

$$B_2 = \frac{\chi_n^*}{\chi_n^{(k-1)}} \tag{S47}$$

which cancels out the last term on the right-hand side of Eq. (S45). Consequently, Eq. (S43) becomes:

$$P_{\text{acc},\chi} = \min\{1, B_0 \cdot B_3\}. \tag{S48}$$

S5.3 GEV parameter $\boldsymbol{\eta}$

The full conditional density for $\boldsymbol{\eta}$ can be determined following derivations similar to those given in Sec. S5.2. From Eq. (S3), we have:

$$p(\boldsymbol{\eta}|\text{rest}) \propto p(\mathbf{Y}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}) \cdot p(\boldsymbol{\eta}|\boldsymbol{\beta}_\eta, \alpha_\eta, \lambda_\eta) \tag{S49}$$

with Eqs. (6) and (S25) leading to:

$$\text{logit}(\boldsymbol{\eta}) \sim \text{MVN}\big(\mathbf{X}_\eta \boldsymbol{\beta}_\eta, \boldsymbol{\Sigma}_\eta(\boldsymbol{\ell}, \alpha_\eta, \lambda_\eta)\big). \tag{S50}$$

Using Eq. (S39) on the transformed variable $\text{logit}(\boldsymbol{\eta})$, the density of $\boldsymbol{\eta}$ can be shown to be:

$$p(\boldsymbol{\eta}|\boldsymbol{\beta}_\eta, \alpha_\eta, \lambda_\eta) = \text{MVN}\big(\text{logit}(\boldsymbol{\eta}); \mathbf{X}_\eta \boldsymbol{\beta}_\eta, \boldsymbol{\Sigma}_\eta\big) \cdot \left(\prod_{s=1}^{S} \eta_s \cdot (1-\eta_s)\right)^{-1}. \tag{S51}$$

MH sampling for the n-th component of $\boldsymbol{\eta}$ is implemented using a logit-normal distribution to generate proposals $\eta_n^* \in [0,1]$:

$$p\left(\eta_n^* | \eta_n^{(k-1)}, \sigma^2_{\text{prop},\eta}\right) = \text{logitN}\left(\eta_n^*; \text{logit}\left(\eta_n^{(k-1)}\right), \sigma^2_{\text{prop},\eta}\right) \tag{S52}$$

leading to the acceptance probability:

$$P_{\text{acc},\eta} = \min\left\{1, \underbrace{\frac{p(\mathbf{Y}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}^*)}{p(\mathbf{Y}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}')}}_{C_0} \cdot \underbrace{\frac{p(\boldsymbol{\eta}^*|\boldsymbol{\beta}_\eta, \alpha_\eta, \lambda_\eta)}{p(\boldsymbol{\eta}'|\boldsymbol{\beta}_\eta, \alpha_\eta, \lambda_\eta)}}_{C_1} \cdot \underbrace{\frac{p\left(\eta_n^{(k-1)}|\eta_n^*, \sigma^2_{\text{prop},\eta}\right)}{p\left(\eta_n^*|\eta_n^{(k-1)}, \sigma^2_{\text{prop},\eta}\right)}}_{C_2}\right\} \tag{S53}$$

with $\boldsymbol{\eta}'$ and $\boldsymbol{\eta}^*$ defined similarly to Eqs. (S29) and (S30), respectively. As per Eq. (S34), the term C_0 becomes:

$$C_0 = \frac{\prod_{t=1}^{T_s}\prod_{i=1}^{D_s} \text{GEV}_r^*\left(\mathbf{y}_{n,t,i}; \tilde{\mu}_n, \sigma_n, \xi_n, \theta_n, \eta_n^*\right)}{\prod_{t=1}^{T_s}\prod_{i=1}^{D_s} \text{GEV}_r^*\left(\mathbf{y}_{n,t,i}; \tilde{\mu}_n, \sigma_n, \xi_n, \theta_n, \eta_n^{(k-1)}\right)} \tag{S54}$$

and analogous to Eq. (S35) with Eq. (S51), the term C_1 simplifies to:

$$C_1 = \underbrace{\exp\left\{\left(\mathbf{X}_\eta \boldsymbol{\beta}_\eta - \text{logit}(\boldsymbol{\eta}'')\right)^T \cdot \left[\boldsymbol{\Sigma}_\eta^{-1}\right]_{(\cdot,n)} \cdot \left(\text{logit}(\eta_n^*) - \text{logit}\left(\eta_n^{(k-1)}\right)\right)\right\}}_{C_3} \times \frac{\eta_n^{(k-1)} \cdot \left(1 - \eta_n^{(k-1)}\right)}{\eta_n^* \cdot (1-\eta_n^*)} \tag{S55}$$

with

$$\boldsymbol{\eta}'' = \left[\eta_1^{(k)}, \ldots, \eta_{n-1}^{(k)}, \eta_n'', \eta_{n+1}^{(k-1)}, \ldots, \eta_S^{(k-1)}\right]^T \tag{S56}$$

$$\eta_n'' = \left(1 + \sqrt{\frac{(1-\eta_n^*)\left(1-\eta_n^{(k-1)}\right)}{\eta_n^* \cdot \eta_n^{(k-1)}}}\right)^{-1}. \tag{S57}$$

Using Eq. (S52), the term C_2 can be shown to simplify as follows:

$$C_2 = \frac{\eta_n^* \cdot (1-\eta_n^*)}{\eta_n^{(k-1)} \cdot \left(1-\eta_n^{(k-1)}\right)} \tag{S58}$$

resulting in:

$$P_{\text{acc},\eta} = \min\{1, C_0 \cdot C_3\}. \tag{S59}$$

S5.4 Regression parameters $\boldsymbol{\beta}_{\tilde{\mu}}$ and $\boldsymbol{\beta}_{\xi}$

For $\chi \in \{\tilde{\mu}, \xi\}$, the full conditional distribution of the respective regression coefficient results from Eqs. (S3), (S7) and (S25) as well as Eq. (6):

$$\begin{aligned} p(\boldsymbol{\beta}_\chi|\text{rest}) &\propto p(\boldsymbol{\chi}|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi) \cdot p(\boldsymbol{\beta}_\chi) && \text{(S60)}\\ &= \text{MVN}\left(\boldsymbol{\chi}; \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi)\right) \cdot \text{MVN}\left(\boldsymbol{\beta}_\chi; \boldsymbol{\mu}_{\beta_\chi}, \boldsymbol{\Sigma}_{\beta_\chi}\right) && \text{(S61)}\\ &= \text{MVN}\left(\boldsymbol{\beta}_\chi; \mathbf{m}_\chi, \mathbf{S}_\chi\right) && \text{(S62)} \end{aligned}$$

where Eq. (S62) results from conjugacy with:

$$\begin{aligned} \mathbf{m}_\chi &= \mathbf{S}_\chi \cdot \left(\boldsymbol{\Sigma}_{\beta_\chi}^{-1}\boldsymbol{\mu}_{\beta_\chi} + \mathbf{X}_\chi^T\boldsymbol{\Sigma}_\chi^{-1}\boldsymbol{\chi}\right) && \text{(S63)}\\ \mathbf{S}_\chi &= \left(\boldsymbol{\Sigma}_{\beta_\chi}^{-1} + \mathbf{X}_\chi^T\boldsymbol{\Sigma}_\chi^{-1}\mathbf{X}_\chi\right)^{-1}. && \text{(S64)} \end{aligned}$$

S5.5 Regression parameters $\boldsymbol{\beta}_{\sigma}$ and $\boldsymbol{\beta}_{\theta}$

Using $\chi \in \{\sigma, \theta\}$, the full conditional distribution of the respective regression coefficient can be derived from Eqs. (S3), (S7) and (S40) as:

$$\begin{aligned} p(\boldsymbol{\beta}_\chi|\text{rest}) &\propto p(\boldsymbol{\chi}|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi) \cdot p(\boldsymbol{\beta}_\chi) && \text{(S65)}\\ &\propto \text{MVN}\left(\log(\boldsymbol{\chi}); \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi)\right) \cdot \text{MVN}\left(\boldsymbol{\beta}_\chi; \boldsymbol{\mu}_{\beta_\chi}, \boldsymbol{\Sigma}_{\beta_\chi}\right) && \text{(S66)}\\ &= \text{MVN}\left(\boldsymbol{\beta}_\chi; \mathbf{m}_\chi, \mathbf{S}_\chi\right) && \text{(S67)} \end{aligned}$$

where Eq. (S67) results from conjugacy with:

$$\begin{aligned} \mathbf{m}_\chi &= \mathbf{S}_\chi \cdot \left(\boldsymbol{\Sigma}_{\beta_\chi}^{-1}\boldsymbol{\mu}_{\beta_\chi} + \mathbf{X}_\chi^T\boldsymbol{\Sigma}_\chi^{-1}\log(\boldsymbol{\chi})\right) && \text{(S68)}\\ \mathbf{S}_\chi &= \left(\boldsymbol{\Sigma}_{\beta_\chi}^{-1} + \mathbf{X}_\chi^T\boldsymbol{\Sigma}_\chi^{-1}\mathbf{X}_\chi\right)^{-1}. && \text{(S69)} \end{aligned}$$

S5.6 Regression parameter $\boldsymbol{\beta}_\eta$

The full conditional distribution for $\boldsymbol{\beta}_\eta$ can be derived from Eqs. (S3), (S7) and (S51) as:

$$p(\boldsymbol{\beta}_\eta|\text{rest}) \propto p(\boldsymbol{\eta}|\boldsymbol{\beta}_\eta, \alpha_\eta, \lambda_\eta) \cdot p(\boldsymbol{\beta}_\eta) \tag{S70}$$

$$\propto \text{MVN}\left(\text{logit}(\boldsymbol{\eta}); \mathbf{X}_\eta\boldsymbol{\beta}_\eta, \boldsymbol{\Sigma}_\eta(\boldsymbol{\ell}, \alpha_\eta, \lambda_\eta)\right) \cdot \text{MVN}\left(\boldsymbol{\beta}_\eta; \boldsymbol{\mu}_{\beta_\eta}, \boldsymbol{\Sigma}_{\beta_\eta}\right) \tag{S71}$$

$$= \text{MVN}\left(\boldsymbol{\beta}_\eta; \mathbf{m}_\eta, \mathbf{S}_\eta\right) \tag{S72}$$

where Eq. (S72) results from conjugacy with:

$$\mathbf{m}_\eta = \mathbf{S}_\eta \cdot \left(\boldsymbol{\Sigma}_{\beta_\eta}^{-1}\boldsymbol{\mu}_{\beta_\eta} + \mathbf{X}_\eta^T\boldsymbol{\Sigma}_\eta^{-1}\,\text{logit}(\boldsymbol{\eta})\right) \tag{S73}$$

$$\mathbf{S}_\eta = \left(\boldsymbol{\Sigma}_{\beta_\eta}^{-1} + \mathbf{X}_\eta^T\boldsymbol{\Sigma}_\eta^{-1}\mathbf{X}_\eta\right)^{-1}. \tag{S74}$$

S5.7 Sill parameters $\alpha_{\tilde{\mu}}$ and α_ξ

Using $\chi \in \{\tilde{\mu}, \xi\}$, the full conditional distribution of the respective sill parameter results from Eqs. (S3), (S6), (S25) and (6) in:

$$p(\alpha_\chi|\text{rest}) \propto p(\boldsymbol{\chi}|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi) \cdot p(\alpha_\chi) \tag{S75}$$

$$= \text{MVN}\left(\boldsymbol{\chi}; \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi)\right) \cdot \text{InvGamma}\left(\alpha_\chi; \kappa_{\alpha_\chi}, \gamma_{\alpha_\chi}\right). \tag{S76}$$

Given the definition of correlation function used in Eq. (7) (exponential correlation), the matrix $\boldsymbol{\Sigma}_\phi(\boldsymbol{\ell}, \alpha_\phi, \lambda_\phi)$ can be decomposed as follows, for $\phi \in \{\tilde{\mu}, \sigma, \xi, \theta, \eta\}$:

$$\boldsymbol{\Sigma}_\phi(\boldsymbol{\ell}, \alpha_\phi, \lambda_\phi) = \alpha_\phi \cdot \tilde{\boldsymbol{\Sigma}}_\phi(\boldsymbol{\ell}, \lambda_\phi) \tag{S77}$$

$$[\tilde{\boldsymbol{\Sigma}}_\phi]_{(i,j)} = \exp\left(\frac{\|\boldsymbol{\ell}_i - \boldsymbol{\ell}_j\|}{\lambda_\phi}\right), \quad i, j = 1, \ldots, S. \tag{S78}$$

Consequently, Eq. (S76) results from conjugacy as:

$$p(\alpha_\chi|\text{rest}) \propto \text{InvGamma}\left(\alpha_\chi; K_\chi, G_\chi\right) \tag{S79}$$

$$K_\chi = \kappa_{\alpha_\chi} + \frac{S}{2} \tag{S80}$$

$$G_\chi = \frac{1}{2} \cdot \left(\boldsymbol{\chi} - \mathbf{X}_\chi\boldsymbol{\beta}_\chi\right)^T \cdot \tilde{\boldsymbol{\Sigma}}_\chi^{-1} \cdot \left(\boldsymbol{\chi} - \mathbf{X}_\chi\boldsymbol{\beta}_\chi\right) + \gamma_{\alpha_\chi}. \tag{S81}$$

S5.8 Sill parameters α_σ and α_θ

For $\chi \in \{\sigma, \theta\}$, the full conditional density of the respective sill parameter can be derived from Eqs. (S3), (S6) and (S40) as:

$$p(\alpha_\chi|\text{rest}) \propto p(\boldsymbol{\chi}|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi) \cdot p(\alpha_\chi) \tag{S82}$$

$$\propto \text{MVN}\left(\log(\boldsymbol{\chi}); \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi)\right) \cdot \text{InvGamma}\left(\alpha_\chi; \kappa_{\alpha_\chi}, \gamma_{\alpha_\chi}\right) \tag{S83}$$

$$= \text{InvGamma}\left(\alpha_\chi; K_\chi, G_\chi\right) \tag{S84}$$

where Eq. (S84) results from conjugacy and Eq. (S77), with:

$$K_\chi = \kappa_{\alpha_\chi} + \frac{S}{2} \tag{S85}$$

$$G_\chi = \frac{1}{2} \cdot \left(\log(\boldsymbol{\chi}) - \mathbf{X}_\chi\boldsymbol{\beta}_\chi\right)^T \cdot \tilde{\boldsymbol{\Sigma}}_\chi^{-1} \cdot \left(\log(\boldsymbol{\chi}) - \mathbf{X}_\chi\boldsymbol{\beta}_\chi\right) + \gamma_{\alpha_\chi}. \tag{S86}$$

S5.9 Sill parameter α_η

The full conditional density for α_η can be derived from Eqs. (S3), (S6) and (S51) as:

$$p(\alpha_\eta|\text{rest}) \propto p(\boldsymbol{\eta}|\boldsymbol{\beta}_\eta, \alpha_\eta, \lambda_\eta) \cdot p(\alpha_\eta) \tag{S87}$$

$$\propto \text{MVN}\left(\text{logit}(\boldsymbol{\eta}); \mathbf{X}_\eta\boldsymbol{\beta}_\eta, \boldsymbol{\Sigma}_\eta(\boldsymbol{\ell}, \alpha_\eta, \lambda_\eta)\right) \cdot \text{InvGamma}\left(\alpha_\eta; \kappa_{\alpha_\eta}, \gamma_{\alpha_\eta}\right) \tag{S88}$$

$$= \text{InvGamma}\left(\alpha_\eta; K_\eta, G_\eta\right) \tag{S89}$$

where Eq. (S89) results from conjugacy and Eq. (S77), with:

$$K_\eta = \kappa_{\alpha_\eta} + \frac{S}{2} \tag{S90}$$

$$G_\eta = \frac{1}{2} \cdot \left(\text{logit}(\boldsymbol{\eta}) - \mathbf{X}_\eta\boldsymbol{\beta}_\eta\right)^T \cdot \tilde{\boldsymbol{\Sigma}}_\eta^{-1} \cdot \left(\text{logit}(\boldsymbol{\eta}) - \mathbf{X}_\eta\boldsymbol{\beta}_\eta\right) + \gamma_{\alpha_\eta}. \tag{S91}$$

S5.10 Range parameters $\lambda_{\tilde{\mu}}$ and λ_ξ

For $\chi \in \{\tilde{\mu}, \xi\}$, the full conditional distribution of the respective range parameter results from Eqs. (S3), (S5), (S25) and (6) as:

$$p(\lambda_\chi|\text{rest}) \propto p(\boldsymbol{\chi}|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi) \cdot p(\lambda_\chi) \tag{S92}$$

$$= \text{MVN}\left(\boldsymbol{\chi}; \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi)\right) \cdot \text{Gamma}\left(\lambda_\chi; \kappa_{\lambda_\chi}, \gamma_{\lambda_\chi}\right) \tag{S93}$$

which represents a non-standard distribution (due to λ_χ being "embedded" within $\boldsymbol{\Sigma}_\chi$), and thus requires a MH sampling step. A non-symmetric proposal distribution is used to generate proposal samples $\lambda_\chi^* > 0$ at the k-th MCMC iteration:

$$p\left(\lambda_\chi^* | \lambda_\chi^{(k-1)}, \sigma^2_{\text{prop},\lambda_\chi}\right) = \text{logN}\left(\lambda_\chi^*; \log\left(\lambda_\chi^{(k-1)}\right), \sigma^2_{\text{prop},\lambda_\chi}\right) \tag{S94}$$

where $\lambda_\chi^{(k-1)}$ is the value of the range parameter at the previous iteration in the MCMC chain. The acceptance probability for the proposal sample is:

$$P_{\text{acc},\lambda_\chi} = \min\left\{1, \frac{\text{MVN}\left(\boldsymbol{\chi}; \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi^*)\right)}{\text{MVN}\left(\boldsymbol{\chi}; \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi^{(k-1)})\right)} \times \underbrace{\frac{\text{Gamma}\left(\lambda_\chi^*; \kappa_{\lambda_\chi}, \gamma_{\lambda_\chi}\right)}{\text{Gamma}\left(\lambda_\chi^{(k-1)}; \kappa_{\lambda_\chi}, \gamma_{\lambda_\chi}\right)} \cdot \frac{p\left(\lambda_\chi^{(k-1)} | \lambda_\chi^*, \sigma^2_{\text{prop},\lambda_\chi}\right)}{p\left(\lambda_\chi^* | \lambda_\chi^{(k-1)}, \sigma^2_{\text{prop},\lambda_\chi}\right)}}_{D_0}\right\} \tag{S95}$$

where D_0 can be shown to simplify to:

$$D_0 = \left(\frac{\lambda_\chi^*}{\lambda_\chi^{(k-1)}}\right)^{\kappa_{\lambda_\chi}} \cdot \exp\left(\frac{\lambda_\chi^{(k-1)} - \lambda_\chi^*}{\gamma_{\lambda_\chi}}\right). \tag{S96}$$

S5.11 Range parameters λ_σ and λ_θ

With $\chi \in \{\sigma, \theta\}$, the full conditional distribution of the range parameters results from Eqs. (S3), (S5) and (S40) as:

$$p(\lambda_\chi|\text{rest}) \propto p(\boldsymbol{\chi}|\boldsymbol{\beta}_\chi, \alpha_\chi, \lambda_\chi) \cdot p(\lambda_\chi) \tag{S97}$$

$$\propto \text{MVN}\left(\log(\boldsymbol{\chi}); \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi)\right) \cdot \text{Gamma}\left(\lambda_\chi; \kappa_{\lambda_\chi}, \gamma_{\lambda_\chi}\right). \tag{S98}$$

Eq. (S98) again represents a non-standard distribution requiring the use of a MH step, which is here also implemented as per Eq. (S94) using a non-symmetric proposal distribution to generate $\lambda_\chi^* > 0$. As per Eqs. (S95) and (S96), the acceptance probability for the proposal sample can be shown to be as follows:

$$P_{\text{acc},\lambda_\chi} = \min\left\{1, \frac{\text{MVN}\left(\log(\boldsymbol{\chi}); \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi^*)\right)}{\text{MVN}\left(\log(\boldsymbol{\chi}); \mathbf{X}_\chi\boldsymbol{\beta}_\chi, \boldsymbol{\Sigma}_\chi\big(\boldsymbol{\ell}, \alpha_\chi, \lambda_\chi^{(k-1)}\big)\right)} \times \left(\frac{\lambda_\chi^*}{\lambda_\chi^{(k-1)}}\right)^{\kappa_{\lambda_\chi}} \cdot \exp\left(\frac{\lambda_\chi^{(k-1)} - \lambda_\chi^*}{\gamma_{\lambda_\chi}}\right)\right\}. \quad \text{(S99)}$$

S5.12 Range parameter λ_η

Finally, the full conditional distribution for λ_η results from Eqs. (S3), (S5) and (S51) as:

$$p(\lambda_\eta|\text{rest}) \propto p(\boldsymbol{\eta}|\boldsymbol{\beta}_\eta, \alpha_\eta, \lambda_\eta) \cdot p(\lambda_\eta) \quad \text{(S100)}$$

$$\propto \text{MVN}\left(\text{logit}(\boldsymbol{\eta}); \mathbf{X}_\eta\boldsymbol{\beta}_\eta, \boldsymbol{\Sigma}_\eta(\boldsymbol{\ell}, \alpha_\eta, \lambda_\eta)\right) \cdot \text{Gamma}\left(\lambda_\eta; \kappa_{\lambda_\eta}, \gamma_{\lambda_\eta}\right). \quad \text{(S101)}$$

As per Eq. (S94), a non-symmetric proposal distribution is used in the MH step for Eq. (S101) so as to generate $\lambda_\eta^* > 0$. And as per Eqs. (S95) and (S96), the acceptance probability for this proposal sample is:

$$P_{\text{acc},\lambda_\eta} = \min\left\{1, \frac{\text{MVN}\left(\text{logit}(\boldsymbol{\eta}); \mathbf{X}_\eta\boldsymbol{\beta}_\eta, \boldsymbol{\Sigma}_\eta(\boldsymbol{\ell}, \alpha_\eta, \lambda_\eta^*)\right)}{\text{MVN}\left(\text{logit}(\boldsymbol{\eta}); \mathbf{X}_\eta\boldsymbol{\beta}_\eta, \boldsymbol{\Sigma}_\eta\big(\boldsymbol{\ell}, \alpha_\eta, \lambda_\eta^{(k-1)}\big)\right)} \times \left(\frac{\lambda_\eta^*}{\lambda_\eta^{(k-1)}}\right)^{\kappa_{\lambda_\eta}} \cdot \exp\left(\frac{\lambda_\eta^{(k-1)} - \lambda_\eta^*}{\gamma_{\lambda_\eta}}\right)\right\}. \quad \text{(S102)}$$

References

Davison A, Padoan S, Ribatet M, 2012. Statistical modelling of spatial extremes. *Statistical Science* **27**(2): 161–186.

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB, 2015. *Bayesian Data Analysis*. Boca Raton: CRC Press, 3rd edition.

Geweke J, 1992. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds.), *Bayesian Statistics 4*, Clarendon Press, Oxford, UK.

Heidelberger P, Welch P, 1983. Simulation run length control in the presence of an initial transient. *Operations Research* **31**: 1109–1144.

Koutsoyiannis D, Kozonis D, Manetas A, 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. *Journal of Hydrology* **206**(1): 118–135.

MacEachern S, Berliner L, 1994. Subsampling the gibbs sampler. *The American Statistician* **48**(3): 188–190.

R Core Team, 2015. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raftery AL, Lewis S, 1992a. How many iterations in the Gibbs sampler? In Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds.), *Bayesian Statistics 4*, Oxford University Press, 763–774.

Raftery AL, Lewis S, 1992b. One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. *Statistical Science* **7**: 493–497.

Spiegelhalter DJ, Best NG, Carlin BP, Linde AVD, 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(4): 583–639.

Stephenson AG, 2016. Bayesian inference for extreme value modelling. In Dey D, Yan J (eds.), *Extreme value modeling and risk analysis: methods and applications*, Chapman & Hall/CRC, 257–280.

Watanabe S, 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research* **11**: 3571–3594.